This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.



us C3>

-CHLAMYDIA PNEUMONIAE GENOME SEQUENCE -

CROSS-REFERENCES TO RELATED APPLICATIONS

The present application is related to 60/128,606, filed April 8, 1999 and 60/108,279, filed November 12, 1998, which are incorporated herein by reference.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

10

FIELD OF THE INVENTION

This invention relates to nucleic acids and polypeptides from *Chlamydia* pneumoniae and to their use in the diagnosis, prevention and treatment of diseases associated with *C. pneumoniae*.

15

20

25

BACKGROUND OF THE INVENTION

Chlamydiaceae is a family of obligate intracellular parasite with a tropism for epithelial cells lining the mucus membranes. The bacteria have two morphologically distinct forms, "elementary body" and "reticulate body". The elementary body is the infectious form, and has a rigid cell wall, primarily of cross-linked outer membrane proteins. The reticulate body is the intracellular, metabolically active form. A unique developmental cycle between these two forms characterizes Chlamydia growth.

C. pneumoniae is a human respiratory pathogen that causes acute respiratory disease, and approximately 10% of community-acquired pneumonia. Antibody prevalence studies have shown that virtually everyone is infected with C. pneumoniae at some time, and that reinfection is common. In addition to respiratory disease, studies have shown an association of this organism with coronary artery disease. It has been demonstrated in atherosclerotic lesions of the aorta and coronary arteries by immunocytochemistry and by polymerase chain reaction (Kuo et al. (1993) J Infect Dis 167(4):841-849).

30

Recent reports have further demonstrated the presence of *C. pneumoniae* in the walls of abdominal aortic aneurysms (Juvonen *et al.* (1997) <u>J Vasc Surg</u> **25**(3):499-505). Abdominal aortic aneurysms are frequently associated with atherosclerosis, and inflammation may be an important factor in aneurysmal dilatation.

10

15

20

25

30



C. pneumoniae may play a role in maintaining an inflammation and triggering the development of aortic aneurysms.

Muhlestein et al. (1996) JACC 27:1555-61, reported a differential incidence of *Chlamydia* species within the coronary artery wall of patients with atherosclerosis versus those with other forms of cardiovascular disease. The extremely high rate of possible infection in patients with symptomatic atherosclerotic disease compared to the very low rate in patients with normal coronary arteries or coronary artery disease from chronic transplant rejection provides evidence for a direct link between the atherosclerotic process and *Chlamydia* infection. Because a history of chlamydial infection is so prevalent in the population, the issue of causality remains. On a physiologic and pathologic level, abnormal interactions among endothelial cells, platelets, macrophages and lymphocytes may lead to a cascade of events resulting in acute endothelial damage, thrombosis and repair, chronically leading to the development of atheroma in blood vessels.

C. pneumoniae is related to other Chlamydia species, but the level of sequence similarity is relatively low. Very little is known about the biology of this organism, although it appears to be an important human pathogen. Allelic diversity and structural relationships between specific genes of Chlamydial species is described in Kaltenboeck et al. (1993) J Bacteriol 175(2):487-502; Gaydos et al. (1992) Infect Immun 60(12):5319-5323; Everett et al. (1997) Int J Syst Bacteriol 47(2):461-473; and Pudjiatmoko et al. (1997) Int J Syst Bacteriol 47(2):425-431.

A number of studies have been published describing methods for detection of *C. pneumoniae*, and for distinguishing between Chlamydial species. Such methods include PCR detection (Rasmussen *et al.* (1992) Mol Cell Probes 6(5):389-394; Holland *et al.* (1990) J Infect Dis 162(4):984-987); a simplified polymerase chain reaction-enzyme immunoassay (Wilson et al. (1996) J Appl Bacteriol 80(4):431-438); sequence determination and restriction endonuclease cleavage (Herrmann *et al.* (1996) J Clin Microbiol 34(8):1897-1902).

Antigenic and molecular analyses of different *C. pneumoniae* strains is described in Jantos *et al.* (1997) <u>J Clin Microbiol</u> 35(3):620-623. Some genes of *C. pneumoniae* have been isolated and sequenced. These include the Gro E operon (Kikuta et al. (1991) <u>Infect Immun</u> 59(12):4665-4669); the major outer membrane protein Perez *et*

al. (1991) Infect Immun 59(6):2195-2199; the DnaK protein homolog (Kornak et al. (1991) Infect Immun 59(2):721-725); as well as a number of ribosomal and other genes.

5

SUMMARY OF THE INVENTION

This invention provides the genomic sequence of *Chlamydia pneumoniae*. The sequence information is useful for a variety of diagnostic and analytical methods. The genomic sequence may be embodied in a variety of media, including computer readable forms, or as a nucleic acid comprising a selected fragment of the sequence. Such fragments generally consist of an open reading frame, transcriptional or translational control elements, or fragments derived therefrom. Proteins encoded by the open reading frames are useful for diagnostic purposes, as well as for their enzymatic or structural activity.

15

20

25

10

DEFINITIONS

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group., e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

30

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

10

15

20

25

30

"Amplification" primers are oligonucleotides comprising either natural or analogue nucleotides that can serve as the basis for the amplification of a select nucleic acid sequence. They include, e.g., polymerase chain reaction primers and ligase chain reaction oligonucleotides.

"Antibody" refers to an immunoglobulin molecule able to bind to a specific epitope on an antigen. Antibodies can be a polyclonal mixture or monoclonal. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoreactive portions of intact immunoglobulins. Antibodies may exist in a variety of forms including, for example, Fv, F_{ab}, and F(ab)₂, as well as in single chains. Single-chain antibodies, in which genes for a heavy chain and a light chain are combined into a single coding sequence, may also be used.

An "antigen" is a molecule that is recognized and bound by an antibody, e.g., peptides, carbohydrates, organic molecules, or more complex molecules such as glycolipids and glycoproteins. The part of the antigen that is the target of antibody binding is an antigenic determinant and a small functional group that corresponds to a single antigenic determinant is called a hapten.

"Biological sample" refers to any sample obtained from a living or dead organism. Examples of biological samples include biological fluids and tissue specimens. Such biological samples can be prepared for analysis of the presence of *C. pneumoniae* nucleic acids, proteins, or antibodies specifically reactive with the proteins.

The term "C. pneumoniae gene" shall be intended to mean the open reading frame encoding specific C. pneumoniae polypeptides, as well as adjacent 5' and 3' non-coding nucleotide sequences involved in the regulation of expression, up to about 2 kb beyond the coding region, but possibly further in either direction. The gene may be introduced into an appropriate vector for extrachromosomal maintenance or for integration into a host genome.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues

H E. 14 į. 90 20

15

5

10

(Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)). Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silen: variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid.

Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

The following groups each contain amino acids that are conservative substitutions for one another:

25 1) Alanine (A), Glycine (G); Serine (S), Threonine (T); 2) 3) Aspartic acid (D), Glutamic acid (E); 4) Asparagine (N), Glutamine (Q); 5) Cysteine (C), Methionine (M); 30 6) Arginine (R), Lysine (K), Histidine (H); 7) Isoleucine (I), Leucine (L), Valine (V); and Phenylalanine (F), Tyrosine (Y), Tryptophan (W). 8) see, e.g., Creighton, Proteins (1984)).

15

20

25

30

10

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. This definition also refers to the complement of a test sequence, which has a designated percent sequence or subsequence complementarity when the test sequence has a designated or substantial identity to a reference sequence. For example, a designated amino acid percent identity of 95% refers to sequences or subsequences that have at least about 95% amino acid identity when aligned for maximum correspondence over a comparison window as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences would then be said to have substantial identity, or to be substantially identical to each other. Preferably, sequences have at least about 70% identity, more preferably 80% identity, more preferably 90-95% identity and above. Preferably, the percent identity exists over a region of the sequence that is at least about 25 amino acids in length, more preferably over a region that is 50-100 amino acids in length.

When percentage of sequence identity is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions, where amino acids residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated according to, e.g., the algorithm of Meyers & Miller, Computer Applic. Biol. Sci. 4:11-17 (1988) e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA)..

15

20

25

30

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated or default program parameters.

A comparison window includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 25 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel et al., supra).

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The

25

30

5

final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, e.g, version 7.0 (Devereaux et al., Nuc. Acids Res. 12:387-395 (1984).

Another example of algorithm that is suitable for determining percent sequence identity (i.e., substantial similarity or identity) is the BLAST algorithm, which 10 is described in Altschul et al., J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues, always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as default parameters a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)).

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below.

Another indication that polynucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence dependent and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically stringent conditions for a Southern blot protocol involve hybridizing in a buffer comprising 5x SSC, 1% SDS at 65°C or hybridizing in a buffer containing 5x SSC and 1% SDS at 42°C and washing at 65°C with a 0.2x SSC, 0.1% SDS wash.

A "label" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ³²P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, dioxigenin, or haptens and proteins for which antisera or monoclonal antibodies are available.

10

15

20

25

30

The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

A labeled nucleic acid probe or oligonucleotide is one that is bound, either covalently, through a linker, or through ionic, van der Waals or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

5

10

15

20

25

30

"Pharmaceutically acceptable" means a material that is not biologically or otherwise undesirable, i.e., the material can be administered to an individual along with a *Chlamydia* antigen without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an analog or mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

The phrase "specifically or selectively hybridizing to," refers to hybridization between a probe and a target sequence in which the probe binds substantially only to the target sequence, forming a hybridization complex, when the target is in a heterogeneous mixture of polynucleotides and other compounds. Such hybridization is determinative of the presence of the target sequence. Although the probe may bind other unrelated sequences, at least 90%, preferably 95% or more of the hybridization complexes formed are with the target sequence.

The term "recombinant" when used with reference to a cell, or nucleic acid, or vector, indicates that the cell, or nucleic acid, or vector, has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

The phrase "specifically immunoreactive with", when referring to a protein or peptide, refers to a binding reaction between the protein and an antibody which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other compounds. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein and are described in detail below.

The phrase "substantially pure" or "isolated" when referring to a Chlamydia peptide or protein, means a chemical composition which is free of other subcellular components of the Chlamydia organism. Typically, a monomeric protein is substantially pure when at least about 85% or more of a sample exhibits a single polypeptide backbone. Minor variants or chemical modifications may typically share the same polypeptide sequence. Depending on the purification procedure, purities of 85%, and preferably over 95% pure are possible. Protein purity or homogeneity may be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band on a polyacrylamide gel upon silver staining. For certain purposes high resolution will be needed and HPLC or a similar means for purification utilized.

DETAILED DESCRIPTION

The present invention provides the nucleotide sequence of the C. pneumoniae genome SEQ ID NO: 1 or a representative fragment thereof, in a form which can be readily used, analyzed, and interpreted by a skilled artisan. As used herein, a "representative fragment" of the nucleotide sequence depicted in SEQ ID NO: 1 refers to any portion which is not presently represented within a publicly available database. Preferred representative fragments of the present invention are open reading frames, expression modulating fragments, uptake modulating fragments, and fragments which can be used to diagnose the presence of C. pneumoniae in sample. Using the information provided in the present application, together with routine cloning and sequencing methods, one of ordinary skill in the art will be able to clone and sequence all "representative fragments" of interest including open reading frames (ORFs) encoding a large variety of C. pneumoniae proteins. A non-limiting identification of such preferred representative fragments is provided in Tables 2 and 2.

Diagnostic use of C. pneumoniae nucleic acids

Hybridization-based assays

Using the nucleic acids disclosed here, one of skill can design nucleic acid hybridization-based assays for the detection of *C. pneumoniae*. Any of a number of well known techniques for the specific detection of target nucleic acids can be used.

Exemplary hybridization-based assays include, but are not limited to, traditional "direct

10

hs

30

20

25

30

5

10

probe" methods such as Southern Blots, dot blots, in situ hybridization (e.g., FISH), PCR, and the like. The methods can be used in a wide variety of formats including, but not limited to substrate- (e.g. membrane or glass) bound methods or array-based approaches as described below. As noted above, this invention also embraces methods for detecting the presence of *Chlamydia* DNA or RNA in biological samples. These sequences can be used to detect *Chlamydia* in biological samples from patients suspected of being infected. A variety of methods of specific DNA and RNA measurement using nucleic acid hybridization techniques are known to those of skill in the art (see Sambrook et al., supra).

In situ hybridization assays are well known (e.g., Angerer (1987) Meth. Enzymol 152: 649). Generally, in situ hybridization comprises the following major steps: (1) fixation of tissue or biological structure to analyzed; (2) prehybridization treatment of the biological structure to increase accessibility of target DNA, and to reduce nonspecific binding; (3) hybridization of the mixture of nucleic acids to the nucleic acid in the biological structure or tissue; (4) post-hybridization washes to remove nucleic acid fragments not bound in the hybridization and (5) detection of the hybridized nucleic acid fragments. The reagent used in each of these steps and the conditions for use vary depending on the particular application.

In a typical in situ hybridization assay, cells are fixed to a solid support, typically a glass slide. If a nucleic acid is to be probed, the cells are typically denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of labeled probes specific to the nucleic acid sequence encoding the protein. The targets (e.g., cells) are then typically washed at a predetermined stringency or at an increasing stringency until an appropriate signal to noise ratio is obtained.

The nucleic acids of this invention are particularly well suited to array-based hybridization formats. Arrays are a multiplicity of different "probe" or "target" nucleic acids (or other compounds) attached to one or more surfaces (e.g., solid, membrane, or gel). In a preferred embodiment, the multiplicity of nucleic acids (or other moieties) is attached to a single contiguous surface or to a multiplicity of surfaces juxtaposed to each other.

In an array format a large number of different hybridization reactions can be run essentially "in parallel." This provides rapid, essentially simultaneous, evaluation

10

15

20

25

30

of a number of hybridizations in a single "experiment". Methods of performing hybridization reactions in array based formats are well known to those of skill in the art (see, e.g., Pastinen (1997) Genome Res. 7: 606-614; Jackson (1996) Nature Biotechnology 14:1685; Chee (1995) Science 274: 610; WO 96/17958.

Arrays, particularly nucleic acid arrays can be produced according to a wide variety of methods well known to those of skill in the art. For example, in a simple embodiment, "low density" arrays can simply be produced by spotting (e.g. by hand using a pipette) different nucleic acids at different locations on a solid support (e.g. a glass surface, a membrane, etc.).

This simple spotting, approach has been automated to produce high density spotted arrays (see, e.g., U.S. Patent No: 5,807,522). This patent describes the use of an automated systems that taps a microcapillary against a surface to deposit a small volume of a biological sample. The process is repeated to generate high density arrays. Arrays can also be produced using oligonucleotide synthesis technology. Thus, for example, U.S. Patent No. 5,143,854 and PCT patent publication Nos. WO 90/15070 and 92/10092 teach the use of light-directed combinatorial synthesis of high density oligonucleotide arrays.

Many methods for immobilizing nucleic acids on a variety of solid surfaces are known in the art. A wide variety of organic and inorganic polymers, as well as other materials, both natural and synthetic, can be employed as the material for the solid surface. Illustrative solid surfaces include, e.g., nitrocellulose, nylon, glass, quartz, diazotized membranes (paper or nylon), silicones, polyformaldehyde, cellulose, and cellulose acetate. In addition, plastics such as polyethylene, polypropylene, polystyrene, and the like can be used. Other materials which may be employed include paper, ceramics, metals, metalloids, semiconductive materials, cermets or the like. In addition, substances that form gels can be used. Such materials include, e.g., proteins (e.g., gelatins), lipopolysaccharides, silicates, agarose and polyacrylamides. Where the solid surface is porous, various pore sizes may be employed depending upon the nature of the system.

In preparing the surface, a plurality of different materials may be employed, particularly as laminates, to obtain various properties. For example, proteins (e.g., bovine serum albumin) or mixtures of macromolecules (e.g., Denhardt's solution) can be employed to avoid non-specific binding, simplify covalent conjugation, enhance

5

10

15

20

25

signal detection or the like. If covalent bonding between a compound and the surface is desired, the surface will usually be polyfunctional or be capable of being polyfunctionalized. Functional groups which may be present on the surface and used for linking can include carboxylic acids, aldehydes, amino groups, cyano groups, ethylenic groups, hydroxyl groups, mercapto groups and the like. The manner of linking a wide variety of compounds to various surfaces is well known and is amply illustrated in the literature.

For example, methods for immobilizing nucleic acids by introduction of various functional groups to the molecules is known (see, e.g., Bischoff (1987) Anal. Biochem., 164: 336-344; Kremsky (1987) Nucl. Acids Res. 15: 2891-2910). Modified nucleotides can be placed on the target using PCR primers containing the modified nucleotide, or by enzymatic end labeling with modified nucleotides. Use of glass or membrane supports (e.g., nitrocellulose, nylon, polypropylene) for the nucleic acid arrays of the invention is advantageous because of well developed technology employing manual and robotic methods of arraying targets at relatively high element densities. Such membranes are generally available and protocols and equipment for hybridization to membranes is well known.

Target elements of various sizes, ranging from 1 mm diameter down to 1 μm can be used. Smaller target elements containing low amounts of concentrated, fixed probe DNA are used for high complexity comparative hybridizations since the total amount of sample available for binding to each target element will be limited. Thus it is advantageous to have small array target elements that contain a small amount of concentrated probe DNA so that the signal that is obtained is highly localized and bright. Such small array target elements are typically used in arrays with densities greater than 10⁴/cm². Relatively simple approaches capable of quantitative fluorescent imaging of 1 cm² areas have been described that permit acquisition of data from a large number of target elements in a single image (see, e.g., Wittrup (1994) Cytometry 16:206-213).

If fluorescently labeled nucleic acid samples are used, arrays on solid surface substrates with much lower fluorescence than membranes, such as glass, quartz, or small beads, can achieve much better sensitivity. Substrates such as glass or fused silica are advantageous in that they provide a very low fluorescence substrate, and a highly efficient hybridization environment. Covalent attachment of the target nucleic acids to glass or synthetic fused silica can be accomplished according to a number of

10

15

20

25

30

known techniques (described above). Nucleic acids can be conveniently coupled to glass using commercially available reagents. For instance, materials for preparation of silanized glass with a number of functional groups are commercially available or can be prepared using standard techniques (see, e.g., Gait (1984) Oligonucleotide Synthesis: A Practical Approach, IRL Press, Wash., D.C.). Quartz cover slips, which have at least 10-fold lower autofluorescence than glass, can also be silanized.

Alternatively, probes can also be immobilized on commercially available coated beads or other surfaces. For instance, biotin end-labeled nucleic acids can be bound to commercially available avidin-coated beads. Streptavidin or anti-digoxigenin antibody can also be attached to silanized glass slides by protein-mediated coupling using e.g., protein A following standard protocols (see, e.g., Smith (1992) Science 258: 1122-1126). Biotin or digoxigenin end-labeled nucleic acids can be prepared according to standard techniques. Hybridization to nucleic acids attached to beads is accomplished by suspending them in the hybridization mix, and then depositing them on the glass substrate for analysis after washing. Alternatively, paramagnetic particles, such as ferric oxide particles, with or without avidin coating, can be used.

A variety of other nucleic acid hybridization formats are known to those skilled in the art. For example, common formats include sandwich assays and competition or displacement assays. Hybridization techniques are generally described in Hames and Higgins (1985) *Nucleic Acid Hybridization, A Practical Approach*, IRL Press; Gall and Pardue (1969) *Proc. Natl. Acad. Sci. USA* 63: 378-383; and John *et al.* (1969) *Nature* 223: 582-587.

Sandwich assays are commercially useful hybridization assays for detecting or isolating nucleic acid sequences. Such assays utilize a "capture" nucleic acid covalently immobilized to a solid support and a labeled "signal" nucleic acid in solution. The sample will provide the target nucleic acid. The "capture" nucleic acid and "signal" nucleic acid probe hybridize with the target nucleic acid to form a "sandwich" hybridization complex. To be most effective, the signal nucleic acid should not hybridize with the capture nucleic acid.

Detection of a hybridization complex may require the binding of a signal generating complex to a duplex of target and probe polynucleotides or nucleic acids.

Typically, such binding occurs through ligand and anti-ligand interactions as between a ligand-conjugated probe and an anti-ligand conjugated with a signal.

10

15

20

25

30

The sensitivity of the hybridization assays may be enhanced through use of a nucleic acid amplification system that multiplies the target nucleic acid being detected. Examples of such systems include the polymerase chain reaction (PCR) system and the ligase chain reaction (LCR) system. Other methods recently described in the art are the nucleic acid sequence based amplification (NASBAO, Cangene, Mississauga, Ontario) and Q Beta Replicase systems.

Nucleic acid hybridization simply involves providing a denatured probe and target nucleic acid under conditions where the probe and its complementary target can form stable hybrid duplexes through complementary base pairing. The nucleic acids that do not form hybrid duplexes are then washed away leaving the hybridized nucleic acids to be detected, typically through detection of an attached detectable label. It is generally recognized that nucleic acids are denatured by increasing the temperature or decreasing the salt concentration of the buffer containing the nucleic acids, or in the addition of chemical agents, or the raising of the pH. Under low stringency conditions (e.g., low temperature and/or high salt and/or high target concentration) hybrid duplexes (e.g., DNA:DNA, RNA:RNA, or RNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (e.g., higher temperature or lower salt) successful hybridization requires fewer mismatches.

One of skill in the art will appreciate that hybridization conditions may be selected to provide any degree of stringency. In a preferred embodiment, hybridization is performed at low stringency to ensure hybridization and then subsequent washes are performed at higher stringency to eliminate mismatched hybrid duplexes. Successive washes may be performed at increasingly higher stringency (e.g., down to as low as 0.25 X SSPE-T at 37°C to 70°C) until a desired level of hybridization specificity is obtained. Stringency can also be increased by addition of agents such as formamide. Hybridization specificity may be evaluated by comparison of hybridization to the test probes with hybridization to the various controls that can be present.

In general, there is a tradeoff between hybridization specificity (stringency) and signal intensity. Thus, in a preferred embodiment, the wash is performed at the highest stringency that produces consistent results and that provides a signal intensity greater than approximately 10% of the background intensity. Thus, in a preferred embodiment, the hybridized array may be washed at successively higher

10

15

20

25

30

stringency solutions and read between each wash. Analysis of the data sets thus produced will reveal a wash stringency above which the hybridization pattern is not appreciably altered and which provides adequate signal for the particular probes of interest.

Methods of optimizing hybridization conditions are well known to those of skill in the art (see, e.g., Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes, Elsevier, N.Y.).

Labeling and detection of nucleic acids.

In a preferred embodiment, the hybridized nucleic acids are detected by detecting one or more labels attached to the sample or probe nucleic acids. The labels may be incorporated by any of a number of means well known to those of skill in the art. Means of attaching labels to nucleic acids include, for example nick translation or endlabeling (e.g. with a labeled RNA) by kinasing of the nucleic acid and subsequent attachment (ligation) of a nucleic acid linker joining the sample nucleic acid to a label (e.g., a fluorophore). A wide variety of linkers for the attachment of labels to nucleic acids are also known. In addition, intercalating dyes and fluorescent nucleotides can also be used.

Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads (*e.g.*, DynabeadsTM), fluorescent dyes (*e.g.*, fluorescein, texas red, rhodamine, green fluorescent protein, and the like, *see, e.g.*, Molecular Probes, Eugene, Oregon, USA), radiolabels (*e.g.*, ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold (*e.g.*, gold particles in the 40 -80 nm diameter size range scatter green light with high efficiency) or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, etc.) beads. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

A fluorescent label is preferred because it provides a very strong signal with low background. It is also optically detectable at high resolution and sensitivity through a quick scanning procedure. The nucleic acid samples can all be labeled with a single label, e.g., a single fluorescent label. Alternatively, in another embodiment, different nucleic acid samples can be simultaneously hybridized where each nucleic acid

5

10

15

20

25

30

sample has a different label. For instance, one target could have a green fluorescent label and a second target could have a red fluorescent label. The scanning step will distinguish cites of binding of the red label from those binding the green fluorescent label. Each nucleic acid sample (target nucleic acid) can be analyzed independently from one another.

Suitable chromogens which can be employed include those molecules and compounds which absorb light in a distinctive range of wavelengths so that a color can be observed or, alternatively, which emit light when irradiated with radiation of a particular wave length or wave length range, e.g., fluorescers.

Desirably, fluorescers should absorb light above about 300 nm, preferably about 350 nm, and more preferably above about 400 nm, usually emitting at wavelengths greater than about 10 nm higher than the wavelength of the light absorbed. It should be noted that the absorption and emission characteristics of the bound dye can differ from the unbound dye. Therefore, when referring to the various wavelength ranges and characteristics of the dyes, it is intended to indicate the dyes as employed and not the dye which is unconjugated and characterized in an arbitrary solvent.

Fluorescers are generally preferred because by irradiating a fluorescer with light, one can obtain a plurality of emissions. Thus, a single label can provide for a plurality of measurable events.

Detectable signal can also be provided by chemiluminescent and bioluminescent sources. Chemiluminescent sources include a compound which becomes electronically excited by a chemical reaction and can then emit light which serves as the detectable signal or donates energy to a fluorescent acceptor. Alternatively, luciferins can be used in conjunction with luciferase or lucigenins to provide bioluminescence. Spin labels are provided by reporter molecules with an unpaired electron spin which can be detected by electron spin resonance (ESR) spectroscopy. Exemplary spin labels include organic free radicals, transitional metal complexes, particularly vanadium, copper, iron, and manganese, and the like. Exemplary spin labels include nitroxide free radicals.

The label may be added to the target (sample) nucleic acid(s) prior to, or after the hybridization. So called "direct labels" are detectable labels that are directly attached to or incorporated into the target (sample) nucleic acid prior to hybridization. In contrast, so called "indirect labels" are joined to the hybrid duplex after hybridization. Often, the indirect label is attached to a binding moiety that has been attached to the

10

15

25

30

target nucleic acid prior to the hybridization. Thus, for example, the target nucleic acid may be biotinylated before the hybridization. After hybridization, an avidin-conjugated fluorophore will bind the biotin bearing hybrid duplexes providing a label that is easily detected. For a detailed review of methods of labeling nucleic acids and detecting labeled hybridized nucleic acids see *Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes*, P. Tijssen, ed. Elsevier, N.Y., (1993)).

Fluorescent labels are easily added during an *in* vitro transcription reaction. Thus, for example, fluorescein labeled UTP and CTP can be incorporated into the RNA produced in an *in vitro* transcription.

The labels can be attached directly or through a linker moiety. In general, the site of label or linker-label attachment is not limited to any specific position. For example, a label may be attached to a nucleoside, nucleotide, or analogue thereof at any position that does not interfere with detection or hybridization as desired. For example, certain Label-ON Reagents from Clontech (Palo Alto, CA) provide for labeling interspersed throughout the phosphate backbone of an oligonucleotide and for terminal labeling at the 3' and 5' ends. As shown for example herein, labels can be attached at positions on the ribose ring or the ribose can be modified and even eliminated as desired. The base moieties of useful labeling reagents can include those that are naturally occurring or modified in a manner that does not interfere with the purpose to which they are put. Modified bases include but are not limited to 7-deaza A and G, 7-deaza-8-aza A and G, and other heterocyclic moieties.

It will be recognized that fluorescent labels are not to be limited to single species organic molecules, but include inorganic molecules, multi-molecular mixtures of organic and/or inorganic molecules, crystals, heteropolymers, and the like. Thus, for example, CdSe-CdS core-shell nanocrystals enclosed in a silica shell can be easily derivatized for coupling to a biological molecule (Bruchez et al. (1998) Science, 281: 2013-2016). Similarly, highly fluorescent quantum dots (zinc sulfide-capped cadmium selenide) have been covalently coupled to biomolecules for use in ultrasensitive biological detection (Warren and Nie (1998) Science, 281: 2016-2018).

Amplification-based assays.

In another embodiment, amplification-based assays can be used to detect nucleic acids. In such amplification-based assays, the nucleic acid sequences act as a

15

20

template in an amplification reaction (e.g. Polymerase Chain Reaction (PCR). Detailed protocols for quantitative PCR are provided in Innis et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, Inc. N.Y.).

Other suitable amplification methods include, but are not limited to ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4: 560, Landegren et al. (1988) Science 241: 1077, and Barringer et al. (1990) Gene 89: 117, transcription amplification (Kwoh et al. (1989) Proc. Natl. Acad. Sci. USA 86: 1173), and self-sustained sequence replication (Guatelli et al. (1990) Proc. Nat. Acad. Sci. USA 87: 1874).

Detection of C. pneumoniae gene expression

The nucleic acids of the invention can also be used to *C. pneumoniae* detect gene transcripts. Methods of detecting and/or quantifying gene transcripts using nucleic acid hybridization techniques are known to those of skill in the art (see Sambrook *et al. supra*). For example, a Northern transfer may be used for the detection of the desired mRNA directly. In brief, the mRNA is isolated from a given cell sample using, for example, an acid guanidinium-phenol-chloroform extraction method. The mRNA is then electrophoresed to separate the mRNA species and the mRNA is transferred from the gel to a nitrocellulose membrane. As with the Southern blots, labeled probes are used to identify and/or quantify the target mRNA.

In another preferred embodiment, the gene transcript can be measured using amplification (e.g. PCR) based methods as described above for directly assessing copy number of the target sequences.

Expression of C. pneumoniae proteins

The nucleic acids disclosed here can be used for recombinant expression

of the proteins. In these methods, the nucleic acids encoding the proteins of interest are introduced into suitable host cells, followed by induction of the cells to produce large amounts of the protein. The invention relies on routine techniques in the field of recombinant genetics, well known to those of ordinary skill in the art. A basic text disclosing the general methods of use in this invention is Sambrook et al., Molecular

Cloning, A Laboratory Manual (2nd ed. 1989).

Standard transfection methods are used to produce prokaryotic, mammalian, yeast or insect cell lines which express large quantities of the desired

10

15

20

25

30

polypeptide, which is then purified using standard techniques (see, e.g., Colley et al., J. Biol. Chem. 264:17619-17622, 1989; Guide to Protein Purification, supra).

The nucleotide sequences used to transfect the host cells can be modified to yield *Chlamydia* polypeptides with a variety of desired properties. For example, the polypeptides can vary from the naturally-occurring sequence at the primary structure level by amino acid, insertions, substitutions, deletions, and the like. These modifications can be used in a number of combinations to produce the final modified protein chain.

The amino acid sequence variants can be prepared with various objectives in mind, including facilitating purification and preparation of the recombinant polypeptide. The modified polypeptides are also useful for modifying plasma half life, improving therapeutic efficacy, and lessening the severity or occurrence of side effects during therapeutic use. The amino acid sequence variants are usually predetermined variants not found in nature but exhibit the same immunogenic activity as naturally occurring protein. In general, modifications of the sequences encoding the polypeptides may be readily accomplished by a variety of well-known techniques, such as site-directed mutagenesis (see Gillman & Smith, Gene 8:81-97 (1979); Roberts et al., Nature 328:731-734 (1987)). One of ordinary skill will appreciate that the effect of many mutations is difficult to predict. Thus, most modifications are evaluated by routine screening in a suitable assay for the desired characteristic. For instance, the effect of various modifications on the ability of the polypeptide to elicit a protective immune response can be easily determined using in vitro assays. For instance, the polypeptides can be tested for their ability to induce lymphoproliferation, T cell cytotoxicity, or cytokine production using standard techniques.

The particular procedure used to introduce the genetic material into the host cell for expression of the polypeptide is not particularly critical. Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasmid vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see Sambrook et al., supra). It is only necessary that the particular procedure utilized be capable of successfully introducing at least one gene into the host cell which is capable of expressing the gene.

5

10

15

20

25

30

Any of a number of well known cells and cell lines can be used to express the polypeptides of the invention. For instance, prokaryotic cells such as *E. coli* can be used. Eukaryotic cells include, yeast, Chinese hamster ovary (CHO) cells, COS cells, and insect cells.

The particular vector used to transport the genetic information into the cell is also not particularly critical. Any of the conventional vectors used for expression of recombinant proteins in prokaryotic and eukaryotic cells may be used. Expression vectors for mammalian cells typically contain regulatory elements from eukaryotic viruses.

The expression vector typically contains a transcription unit or expression cassette that contains all the elements required for the expression of the polypeptide DNA in the host cells. A typical expression cassette contains a promoter operably linked to the DNA sequence encoding a polypeptide and signals required for efficient polyadenylation of the transcript. The term "operably linked" as used herein refers to linkage of a promoter upstream from a DNA sequence such that the promoter mediates transcription of the DNA sequence. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

Following the growth of the recombinant cells and expression of the polypeptide, the culture medium is harvested for purification of the secreted protein. The media are typically clarified by centrifugation or filtration to remove cells and cell debris and the proteins are concentrated by adsorption to any suitable resin or by use of ammonium sulfate fractionation, polyethylene glycol precipitation, or by ultrafiltration. Other routine means known in the art may be equally suitable. Further purification of the polypeptide can be accomplished by standard techniques, for example, affinity chromatography, ion exchange chromatography, sizing chromatography, His₆ tagging and Ni-agarose chromatography (as described in Dobeli *et al.*, *Mol. and Biochem. Parasit.* 41:259-268 (1990)), or other protein purification techniques to obtain homogeneity. The purified proteins are then used to produce pharmaceutical compositions, as described below.

An alternative method of preparing recombinant polypeptides useful as vaccines involves the use of recombinant viruses (e.g., vaccinia). Vaccinia virus is grown

5

10

15

20

25

30

in suitable cultured mammalian cells such as the HeLa S3 spinner cells, as described by Mackett et al., in DNA cloning Vol. II: A practical approach, pp. 191-211 (Glover, ed.).

Antibody Production

The proteins of the present invention can be used to produce antibodies specifically reactive with *C pneumoniae* antigens. If isolated proteins are used, they may be recombinantly produced or isolated from *Chlamydia* cultures. Synthetic peptides made using the protein sequences may also be used.

Methods of production of polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera is prepared. Further fractionation of the antisera to enrich for antibodies reactive to *Chlamydia* proteins can be done if desired (see Harlow & Lane, *Antibodies: A Laboratory Manual* (1988)).

Polyclonal antisera are used to identify and characterize *Chlamydia* in the tissues of patients using, for instance, *in situ* techniques and immunoperoxidase test procedures described in Anderson *et al. JAVMA* 198:241 (1991) and Barr *et al. Vet. Pathol.* 28:110-116 (1991).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see Kohler & Milstein, Eur. J. Immunol. 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

Monoclonal-antibodies produced in such a manner are used, for instance, in ELISA diagnostic tests, immunoperoxidase tests, immunohistochemical tests, for the *in vitro* evaluation of spirochete invasion, to select candidate antigens for vaccine development, protein isolation, and for screening genomic and cDNA libraries to select appropriate gene sequences.

5

10

15

20

25

30

Immunodiagonostic detection of C. pneumoniae infections

The present invention also provides methods for detecting the presence or absence of *C. pneumoniae*, or antibodies reactive with it, in a biological sample. For instance, antibodies specifically reactive with *Chlamydia* can be detected *using* either *Chlamydia* proteins or the isolates described here. The proteins and isolates can also be used to raise specific antibodies (either monoclonal or polyclonal) to detect the antigen in a sample. In addition, the nucleic acids disclosed and claimed here can be used to detect *Chlamydia*-specific sequences using standard hybridization techniques.

For a review of immunological and immunoassay procedures in general, see *Basic and Clinical Immunology* (Stites & Terr ed., 7th ed. 1991)). The immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay* (Maggio, ed., 1980); Tijssen, *Laboratory Techniques in Biochem stry and Molecular Biology* (1985)). For instance, the proteins and antibodies disclose I here are conveniently used in ELISA, immunoblot analysis and agglutination assays.

In brief, immunoassays to measure anti-Chlamydia antibodies or antigens can be either competitive or noncompetitive binding assays. In competitive binding assays, the sample analyte (e.g., anti-Chlamydia antibodies) competes with a labeled analyte (e.g., anti-Chlamydia monoclonal antibody) for specific binding sites on a capture agent (e.g., isolated Chlamydia protein) bound to a solid surface. The concentration of labeled analyte bound to the capture agent is inversely proportional to the amount of free analyte present in the sample.

Noncompetitive assays are typically sandwich assays, in which the sample analyte is bound between two analyte-specific binding reagents. One of the binding agents is used as a capture agent and is bound to a solid surface. The second binding agent is labelled and is used to measure or detect the resultant complex by visual or instrument means.

A number of combinations of capture agent and labelled binding agent can be used. For instance, an isolated *Chlamydia* protein or culture can be used as the capture agent and labelled anti-human antibodies specific for the constant region of human antibodies can be used as the labelled binding agent. Goat, sheep and other non-human antibodies specific for human immunoglobulin constant regions (e.g., γ or μ) are

5

10

15

20

25

30

well known in the art. Alternatively, the anti-human antibodies can be the capture agent and the antigen can be labelled.

Various components of the assay, including the antigen, anti-Chlamydia antibody, or anti-human antibody, may be bound to a solid surface. Many methods for immobilizing biomolecules to a variety of solid surfaces are known in the art. For instance, the solid surface may be a membrane (e.g., nitrocellulose), a microtiter dish (e.g., PVC or polystyrene) or a bead. The desired component may be covalently bound or noncovalently attached through nonspecific bonding.

Alternatively, the immunoassay may be carried out in liquid phase and a variety of separation methods may be employed to separate the bound labeled component from the unbound labelled components. These methods are known to those of skill in the art and include immunoprecipitation, column chromatography, adsorption, addition of magnetizable particles coated with a binding agent and other similar procedures.

An immunoassay may also be carried out in liquid phase without a separation procedure. Various homogeneous immunoassay methods are now being applied to immunoassays for protein analytes. In these methods, the binding of the binding agent to the analyte causes a change in the signal emitted by the label, so that binding may be measured without separating the bound from the unbound labelled component.

Western blot (immunoblot) analysis can also be used to detect the presence of antibodies to *Chlamydia* in the sample. This technique is a reliable method for confirming the presence of antibodies against a particular protein in the sample. The technique generally comprises separating proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the separated proteins. This causes specific target antibodies present in the sample to bind their respective proteins. Target antibodies are then detected using labeled antihuman antibodies.

The immunoassay formats described above employ labelled assay components. The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. A wide variety of labels may be used. The component may be labelled by any one of several methods. Traditionally a radioactive label incorporating ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P was used. Non-radioactive labels

include ligands which bind to labelled antibodies, fluorophores, chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labelled ligand. The choice of label depends on sensitivity required, ease of conjugation with the compound, stability requirements, and available instrumentation.

Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labelling or signal producing systems which may be used, see U.S. Patent No. 4,391,904, which is incorporated herein by reference.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to an anti-ligand (e.g., streptavidin) molecule which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. A number of ligands and anti-ligands can be used. Where a ligand has a natural anti-ligand, for example, biotin, thyroxine, and cortisol, it can be used in conjunction with the labelled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

Some assay formats do not require the use of labelled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labelled and the presence of the target antibody is detected by simple visual inspection.

Pharmaceutical Compositions

5

10

15

20

25

30

M

The peptides or antibodies (typically monoclonal antibodies) of the present invention and pharmaceutical compositions thereof are useful for administration to mammals, particularly humans, to treat and/or prevent *Chlamydia* infections. Suitable formulations are found in *Remington's Pharmaceutical Sciences*, Mack Publishing Company, Philadelphia, PA, 17th ed. (1985).

10

15

20

25

30

The immunogenic peptides or antibodies of the invention are administered prophylactically or to an individual already suffering from the disease. The peptide compositions are administered to a patient in an amount sufficient to elicit an effective immune response to *Chlamydia*. An effective immune response is one that inhibits infection. An amount adequate to accomplish this is defined as "therapeutically effective dose" or "immunogenically effective dose." Amounts effective for this use will depend on, e.g., the peptide composition, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician, but generally range for the initial immunization (that is for therapeutic or prophylactic administration) from about 0.1 mg to about 1.0 mg per 70 kilogram patient, more commonly from about 0.5 mg to about 0.75 mg per 70 kg of body weight. Boosting dosages are typically from about 0.1 mg to about 0.5 mg of peptide using a boosting regimen over weeks to months depending upon the patient's response and condition. A suitable protocol would include injection at time 0, 4, 2, 6, 10 and 14 weeks, followed by further booster injections at 24 and 28 weeks.

For therapeutic use, administration should begin at the first sign of infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In some circumstances, loading doses followed by boosting doses may be required. The resulting immune response helps to cure or at least partially arrest symptoms and/or complications. Vaccine compositions containing the peptides are administered prophylactically to a patient susceptible to or otherwise at risk of the infection.

The pharmaceutical compositions (containing either peptides or antibodies) are intended for parenteral or oral administration. Preferably, the pharmaceutical compositions are administered parenterally, e.g., subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic polypeptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, e.g., water, buffered water, 0.4% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain

10

. 15

20

25

pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

The compositions may also comprise carriers to enhance the immune response. Useful carriers are well known in the art, and include, e.g., KLH, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly(lysine:glutamic acid), influenza, hepatitis B virus core protein, hepatitis B virus recombinant vaccine and the like.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed ty incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

As noted above, the peptide compositions are intended to induce an immune response to *Chlamydia*. Thus, compositions and methods of administration suitable for maximizing the immune response are preferred. For instance, peptides may be introduced into a host, including humans, linked to a carrier or as a homopolymer or heteropolymer of active peptide units from various *Chlamydia* proteins disclosed here. Alternatively, a "cocktail" of polypeptides can be used. A mixture of more than one polypeptide has the advantage of increased immunological reaction and, where different peptides are used to make up the polymer, the additional ability to induce antibodies to a number of epitopes.

The compositions also include an adjuvant. As used here, number of adjuvants are well known to one skilled in the art. Suitable adjuvants include incomplete Freund's adjuvant, alum, aluminum phosphate, aluminum hydroxide,

N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP),

N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-Lalanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria, monophosphoryl

20

25

lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. The effectiveness of an adjuvant may be determined by measuring the amount of antibodies directed against the immunogenic peptide.

The concentration of immunogenic peptides of the invention in the pharmaceutical formulations can vary widely, i.e. from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, etc., in accordance with the particular mode of administration selected.

The peptides of the invention can also be expressed by attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (*Nature* 351:456-460 (1991)). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g., *Salmonella typhi* vectors and the like, will be apparent to those skilled in the art from the description herein.

The DNA encoding one or more of the peptides of the invention can also be administered to the patient. This approach is described, for instance, in Wolff et. al., Science 247: 1465-1468 (1990) as well as U.S. Patent Nos. 5,580,859 and 5,589,466.

In order to enhance serum half-life, the peptides may also be encapsulated, introduced into the lumen of liposomes, prepared as a colloid, or other conventional techniques may be employed which provide an extended serum half-life of the peptides. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al., *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), U.S. Pat. Nos. 4, 235,871, 4,501,728 and 4,837,028.

30

EXAMPLES

The following examples are offered to illustrate, but no to limit the claimed invention.

Example 1:

10

15

20

25

30

This example describes comparison of the *C. pneumoniae* genome disclosed here and the, previously sequenced, *C. trachomatis* genome (Stephens, *et al. Science* **282:**754-759 (1998)).

The apparent low level of DNA homology between *C. trachomatis* and *C. pneumoniae* (Campbell, *et al.*, *J. Clin. Microbiol.* **25**:1911-1916 (1987)) yet analogous cell structures and developmental cycles, predicts that comparative analysis of the two genomes will significantly enhance the understanding of both pathogens. Identification of genes that are present in one species but not the other are of particular importance for the mutually exclusive biological, virulence and pathogenesis capabilities of each.

Identification of genes shared between the two species strongly supports the requirement for these capabilities in a biological system that has, over its long-term association with mammalian host cells, evolved to reduce the metabolic capacities while optimizing survival, growth and transmission of these unique pathogens.

The previously sequenced *C. trachomatis* genome contains 1,042,519 nucleotides and 875 likely protein-coding genes. Similarity searching permitted the inferred functional assignment of sequences 636 (60%) genes disclosed here and 251 (23%) are similar to hypothetical genes for other bacterial organisms including those for *C. trachomatis*. The remaining 186 (17%) genes are not homologous to sequences deposited in GenBank. Seventy *C. trachomatis* genes are not represented in the *C. pneumoniae* genome. These are contained within blocks consisting of 2-17 genes and 19 single genes. Of the 70 *C. trachomatis* genes without homologs in *C. pneumoniae*, 60 are classified as encoding hypothetical proteins. The remaining genes not represented in *C. pneumoniae* consist of the tryptophan operon (*trpA,B,R*), *trpC*, two predicted thiol protease genes, and 4 genes assigned to the phospholipase-D superfamily.

It is evident that there is a high level of functional conservation between C. pneumoniae and C. trachomatis as orthologs to C. trachomatis genes were identified for 859 (80%) of the predicted coding sequences for C. pneumoniae. The level of similarity for individual encoded proteins spans a wide spectrum (22-95% amino acid identity) with an average of 62% amino acid identity between orthologs from the two species. The percent amino acid identity between orthologous chlamydial proteins is similar among functional groups with the highest for proteins associated with translation and the lowest for proteins whose function in chlamydiae is uncharacterized and not related to proteins encoded by other organisms. The gene order of the homologous set of genes in C.

10

15

20

25

30

pneumoniae shows reorganization relative to the genome of *C. trachomatis*; however, there is a high level of synteny for the gene organization of the two genomes. We identified thirty-nine blocks of 2 or more genes whose gene organization is colinear with homologs to *C. trachomatis*, although some of these are inverted. The distribution of genome reorganization is not evenly distributed on the chromosome as the region between *C. pneumoniae* coding sequences 0130-0300 contains substantially more reorganization than other areas of the genome. This region coincides with the predicted chromosome replication terminus.

We identified orthologs of enzymes characterized in other bacteria that account for the essential requirements for DNA replication, repair, transcription and translation including two predicted DNA helicases of the Swi2/Snf2 family found in *C. trachomatis*. Similar to *C. trachomatis*, alternative sigma subunits for RNA polymerase, σ²⁸ and σ⁵⁴, were identified in addition to anti-σ regulatory system factors RsbV, a RsbW-like single-domain histidine kinase, and a RsbU-like protein phosphatase. These findings suggest that the fundamental mechanisms of transcriptional regulation are conserved among *Chlamydia*. The *C. trachomatis* proteins containing SET and SWIB domains, and a SWIB domain fused to the C-terminus of the chlamydial topoisomerase I, not identified outside eukaryotes, are found in *C. pneumoniae* supporting their possible role in the chromatin condensation-decondensation characteristic of the biologically unique chlamydial developmental cycle.

The central metabolic pathways inferred from the *C. pneumoniae* genome sequence are the same as those identified for *C. trachomatis C. pneumoniae* has a glycolytic pathway and a linked tricarboxylic acid cycle, although likely functional, is incomplete as genes for citrate synthase, aconitase, and isocitrate dehydrogenase were not identified. *C. pneumoniae* has a complete glycogen synthesis and degradation system supporting a role for glycogen synthesis and utilization of glucose-derivatives in chlamydial metabolism. Genes encoding essential functions in aerobic respiration are present and electron flux may be supported by pyruvate, succinate, glycerol-3-phosphate, and NADH dehydrogenases, NADH-ubiquinone oxidoreductase and cytochrome oxidase. *C. pneumoniae* also contains the V (vacuolar)-type ATPase operon and the two ATP translocases found in *C. trachomatis*.

The type-III secretion virulence system required for invasion by several pathogenic bacteria and found in the *C. trachomatis* genome in three chromosomal

15

20

25

30

5

10

locationsis also present in the *C. pneumoniae* genome. Each of the components is conserved and their relative genomic contexts are conserved. Genes such as a predicted serine/threonine protein kinase and other genes physically linked to genes encoding structural components of the type-III secretion apparatus, but without identified homologs, are also highly similar between the two species suggesting the functional roles in modifying cellular biology are fundamentally conserved.

Chlamydia-encoded proteins that are not found in chlamydial organisms but localized to the intracellular chlamydial inclusion membrane are likely essential for the unique intracellular biology and perhaps differences in inclusion morphology observed between species of Chlamydia. Several such proteins, termed IncA,B&C, have been characterized for a C. psittaci strain (Rockey, et al. Mol. Microbiol. 15:617-626 (1995); Rockey et al. Infect. Immun. 62:106-112 (1994)). C. pneumoniae and C. trachomatis encode orthologs to C. psittaci IncB and IncC and C. trachomatis also contains an ortholog to IncA. C. pneumoniae contains two genes that encode proteins with similarity to IncA (CPn0186 and CPn0585), although the level of homology is low suggesting analogous but possibily altered functions.

The tryptophan biosynthesis operon (trpA, trpB, trpR) and trpC identified in C. trachomatis is conspicuously missing in the C. pneumoniae genome. This represents the entire repertoire of genes associated with tryptophan biosynthesis identified in C. trachomatis. Seventeen genes adjacent to the C. trachomatis tryptophan operon also were not found in the C. pneumoniae genome. This region is the single largest loss of a contiguous genomic segment and includes 4 HKD superfamily encoding genes that encompass a family of proteins related to endonuclease and phospholipase D. These findings may be important for the ability of Chlamydia to persist in their hosts and cause disease by eliciting potent, focal and persistent inflammatory responses thought to be essential for pathogenesis.

The *C. pneumoniae* genome contains 187,711 additional nucleotides compared to the *C. trachomatis* genome, and the 214 coding sequences not found in *C. trachomatis* account for most of the increased genome size. Eighty-eight of these genes are found in blocks of >10 genes (11-30 genes/block), 41 are single genes, and the remainder are partnered with at least one other gene. Based upon the observation that ~70% of all the *C. pneumoniae* genes have an identifiable homolog in GenBank, exclusive of *C. trachomatis*, it would be expected that over 150 of the 214 genes should

10

15

25

30

have a homolog in GenBank, many associated with a function. However, only 28 coding sequences have similarity to genes from other organisms. Thus the majority of the genes that are mutually exclusive of C. trachomatis (186 of 214), and the 60 of 70 C. trachomatis genes that lacked an identifiable homolog in C. pneumoniae, do not have detectable homologs to genes from other organisms. We predict that most of the unique genes are essential for specific attributes that define the differential biology, tropism and pathogenesis of C. trachomatis and C. pneumoniae. Moreover, this suggests that C. pneumoniae has more unique biological (i.e., virulence) capacity than C. trachomatis. The ability of C. pneumoniae to be more invasive and survive in a broader range of host cell types than C. trachomatis is consistent with this hypothesis. Not all of the differences in biological capacity may be associated with mutually exclusive genes. One explanation for the significantly lower level of homology between protein sequences assigned as having C. pneumoniae and C. trachomatis orthologs but no identifiable orthologs in other organisms is that this set of proteins is not only associated with biological requirements specific for Chlamydia but this polymorphism may account for differential biology between the two species. The determination of the genome sequence from a representative of the C. psittaci group will precisely delineate those genes that are mutually exclusive and specific for each species.

The major functionally identifiable addition to the *C. pneumoniae* genome is a large expansion of genes encoding a new family of chlamydial polymorphic membrane proteins (Pmp), alone representing 22% of the increased coding capacity. While the *C. trachomatis* genome has 9 *pmp* genes, remarkably the *C. pneumoniae* genome contains 21 *pmp* genes. Most of these genes appear to be amplified in two regions of the genome with three stand-alone genes. Interestingly one of the stand-alone genes is most closely related to the *C. trachomatis pmpD* which is the only stand-alone *pmp* gene in the *C. trachomatis* genome and it is located with the same relative genomic context, suggesting an essential and conserved function for this paralog. Six Pmp-coding genes are presumably not functional as five contain predicted coding frame-shifts and one is truncated. The amplification of this gene family and the confidently predicted frame-shifts suggest a specific molecular mechanism to promote functional or antigenic diversity. The biological role of this protein family remains enigmatic, although at least one of the proteins in *C. psittaci* related to this family is exposed on the chlamydial surface.

15

20

25

While a function could not be assigned for most of the unique *C. pneumoniae* genes, several have significant similarity to genes from other organisms. Functional assignments could be made for genes encoding GMP synthetase, IMP dehydrogenase, UMP synthase, uridine kinase, biotin synthase pathway proteins, methylthioadenosine nucleosidase, a DNA glycosylase and aromatic amino acid hydroxylase. Thus a complete pathway was identified for biotin biosynthesis. The additional purine and pyrimidine salvage pathway genes presumably reflect metabolic limitations in one of the cell types that *C. pneumoniae* infects or differences in the ability of *C. pneumoniae* to transport precursor nucleosides or nucleotides.

The addition of aromatic amino acid hydroxylase in *C. pneumoniae* is intriguing especially in light of the loss of tryptophan biosynthetic genes and the inability to synthesize other amino acids including phenylalanine. Aromatic amino acid hyroxlyases include three distinct enzymes that function to receptively oxidize phenylalanine to tyrosine, tyrosine to Dopa, and tryptophan to 5-hydroxytryptophan and serotonin. Although the chlamydial protein is similar to proteins of this family and incrementally more closely related to tryptophan hydroxylase, its specific function could not be confidently predicted. We hypothesize that it may be involved in *C. pneumoniae* virulence. Tryptophan hydroxylase has not been previously identified in bacteria and the origin of the chlamydial gene appears to be from eukaryotes. The functional role of an aromatic amino acid hydroxylase for *C. pneumoniae* is linked to the unique intracellular biology of this organism and may represent a key contribution to *C. pneumoniae* persistence and pathogenesis.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

30

Table 1 provides functional assignments of *C. pneumoniae* nonproteinencoding genomic sequences. Table 2 provides functional assignments of protein coding sequences. Table 3 provides the amino acid sequences of the proteins corresponding to the coding sequences.

TABLE 1

ty	pe e	SEQ ID NO:1 start position	SEQ ID No end position		Gene		
	O D D R R NA A A A A A A A A A A A A A A A A	start position 841664 138493 607342 1000564 1002415 1005393 269070 164318 296224 836191 1030533 784896 781680 961536 999949 268992 672236 680178 715889 739403 1175863 784994 843926 409922 631373 677337 807413 877473 462141 1085605 786780 89728 293477 87522 199301 199390 626904 708359 1142034 1230028 91070	end position 841396 138074 607649 1005278 10055278 10055278 10055278 1005542 164389 296151 836119 1030603 784822 781610 961607 1000023 269065 672318 680257 715971 739486 1175944 784922 843999 409848 631445 677264 807341 877400 462214 1085676 786708 89657 293405 87450 199229 199317 626987 708440 1142117 1229945 90999	(Ribss an) yseuuuuu) o) e)) gn MM MtttsT (Ribss an) yseuuuuu) o) e))) gn)))) yseuuuuu) o) e))) gn)))) yseeri)	RNA RNA RNA RNA RNA RNA RNA RNA RNA Glu tRNA Glu tRNA RNA RNA RNA RNA RNA RNA RNA RNA RNA	P RNA	Replica
	ERNA ERNA ERNA	293399 296147 1137389	293317 296075 1137462		yr tRNA al tRNA_1 RNA_2		

TABLE 2

```
in parentheses!
                                        Gene Function (C. trachomatis ortholog
                              Strand
Gene_1
                                        CT001 hypothetical protein
                                        gatC-Glu-tRNA Gln Amidotransferase (C subunit) - (CT002)
                                  R
           282
CPn0001
                                  •
                       875
CPn0002
           573
                                        gatA-Glu tRNA Gln Amidotransferae-(CT003)
                                        gatB (Petil2) Glu FPHA Gln Amidotransferase (B Subunit) - (CT904)
                                  P
                       2370
           895
CPn0003
                       : (: 1
                                  Į?
                                        pmp_i-Polymorphic Outer Membrane Protein G Family
           2170
 rn0004
                                  F
                       6832
           4127
CPROUUS
                                  R
                       7141
           7293
CPn0006
                       10496
CPn0007
           7605
                       11685
                                  F
           10975
CPn0008
                       13119
           11815
CPn0009
                                  P
                       14325
CPn0010
           13435
                                        frame-shift with 0010
                       15746
                                  P
           14379
CPn0010
                       16614
CPn0011
           15892
                                        pmp_2-Polymorphic Outer Membrane Protein G Family
                       18212
           16644
CPn0012
                                        pmp_3-Polymorphic Outer Membrane Protein G Family
                                  F
                       21106
           18584
CPn0013
                       21922
                                  F
           21392
                                        pmp_3-PMP_3 (frame-shift with 0014)
CPn0014
                       24174
                                  F.
                                        pmp_4-Polymorphic Outer Membrane Protein G Family
           21835
CPn0015
                                  £,
                       26188
           24416
                                        pmp_4-PMP_4 (frame-shift with 0016)
CPn0016
                       27170
                                  £.
                                        pmp_5-Polymorphic Outer Membrane Protein G Family
           26094
CPn0017
                                  F
                       29003
           27522
                                        pmp_5-PMP_5 (frame-shift with 0018)
CP#0018
                                        Predicted OMP [leader (14) peptide: outer membrane]-(CT351)
                       30356
                                  F
           29007
CPn0019
                       30603
                                        Predicted OMP [leader (19) peptide]-(CT350)
           32687
CPn0020
                       32707
           34410
CPn0021
                                        maf-(CT349)
                                        yjjK/alr-ABC Transporter Protein ATPase-(CT348)
                       34395
CPh0022
           34982
                       35014
                                  R
           36603
CP#0023
                                        xerC-Integrase/recombinase-(CT347)
                                        elaC/atsA-Sulphohydrolase/Glycosulfatase-( CT346)
CPn0024
                       36661
            37596
                       37684
                                  R
            38604
                                        CT345 hypothetical protein-(CT345)
CP50025
                                  R
                       38762
            39625
                                        lon-Lon ATP-dependent Protease-(CT344)
CP#0026
                                   R
                       39778
CPn0027
            42234
                                   R
                       42543
            43325
CPp0028
                       43390
                                   R
                                        gcp_1-0-Sialoglycoprotein Endopeptidase_1-(CT343)
            43755
CPn0029
                                   F
                       44529
C5P0030
            43891
                                        rs21-S21 Ribosomal Protein-(CT342)
                       44884
CP#0031
            44711
                                        dnaJ-Heat Shock Protein J-(CT341)
                                        pdhA&B/odbAkodbB-(pyruvate) Oxoisovalerate Dehydrogenase Alpha & Beta
                                   P
                       46098
            44923
CP#0032
                       48171
                                   P
            46138
СР#2033
                                           Fusion-(CT340)
 CPn0034
                       48210
                                   R
            49457
                                        CT339 hypothetical protein
                       49569
                                   R
 CP00035
            51029
                                        CT338 hypothetical protein
                                   F
                       51796
                                        ptsH-PTS Phosphocarrier Protein Hpr-(CT337)
 CPn0036
            51002
                                   F
            51792
                        52115
                                         ptsI-PTS PEP Phosphotransferase-(CT336)
 CPn0037
                                   F
                        53831
 CPn0038
            52119
                                        ybaB-(CT335)
                        53963
                                   R
            54250
 EPn0039
                                         dnax_1-DNA Pol III Gamma and Tau_1-(CT334)
                        54318
            55643
 CPn0040
            55996
                        57342
                                   P
 CPn0041
                        58182
 CPn0042
            57403
            58447
                        60372
                                   F
 CPn0043
                        60778
            60419
 CPn0044
                                   F
            61069
                        62790
 CPn0045
                        63263
                                   F
            62790
 CPn0046
                                         *YqfF-Bs conserved hypothetical IM protein
 CPn0047
            63455
                        63652
                                   F
                        65801
            63687
 CPn0048
                                   R
            66296
                        65817
 CPn0049
                        66499
                                   R
            66813
 CPn0050
             66833
                        67111
                                   P
 CPn0051
                                         hemC-Porphobilinogen Deaminase-(CT299)
            68005
                        67304
---CPn0052-
                                         sms-Sms Protein-(CT298)
             69344
                        67986
                                   R
 CPn0053
                                         rnc-Ribonuclease III-(CT297)
             70023
                        69313
                                   R
 CPn0054
                                         CT296 hypothetical protein
             70129
                        70590
                                   F
 CPn0055
                                         mrsA-Phosphomannomutase-(CT295)
             70953
                        72746
                                         sodM-Superoxide Dismutase (Mn)-(CT294)
 CPn0056
                                         accD-AcCoA Carboxylase/Transferase Beta-(CT293)
                        73554
                                    F
 CPn0057
             72934
                        74562
             73639
 CPn0058
                                         dut-dUTP Nucleotidohydrolase-(CT292)
                        75050
             14616
 CPn0059
                                         ptsN_1-PTS IIA Protein-(CT291)
                                         ptsN_2-PTS IIA Protein + HTH DNA-Binding Domain-(CT290)
                        75528
 CPn0060
             75055
                        76208
                                    F
 CPn0061
             75534
                                         CT289 hypothetical protein
                        77690
 CPn0062
             76308
                        78267
 CPn0063
             78112
                        78576
             74346
 CPR0054
                                         CT288 hypothetical protein
                        8065 L
                                    F
 CPn0065
             78924
                         82655
             H0925
 CPROOSS
```

```
84053
             82953
  CPn0067
                                          CT360 hypothetical protein
                                    R
             84903
                         84331
  CPn0068
                         87086
  CPn0069
             85236
                         87208
                                    R
             87378
  CPn0070
                                          CT325 hypothetical protein
                         87599
                                    R
             88045
  CPn0071
                                          CT324 hypothetical protein
                         88057
                                    R
  CPn0072
             89061
                                          infA-Initiation Factor IF-1-(CT323)
                                    F
                         89574
             89356
  CPn0073 '
                                          tufA-Elongation Factor Tu-(CT322)
                         90955
                                    F
             89774
  CPn0074
                                          secE-preprotein translocase-(CT321)
                         91350
             91102
  CPn0075
                                          nusG-Transcriptional Antitermination-(CT320)
                         91903
  CPn0076
             91358
                                          rlll-Lll Ribosomal Protein-(CT319)
                                    F
                         92435
             92013
  CPn0077
                                          rll-L1 Ribosomal Protein-(CT318)
                         93160
             92465
  CPn0078
                                          rll0-Ll0 Ribosomal Protein-(CT317)
                         93688
  CPn0079
             93179
                                          rl7-L7/L12 Ribosomal Protein-(CT316)
                         94121
             93735
  CPn0080
                                          rpoB-RNA Polymerase Beta-(CT315)
             94261
                         98016
                                    F
  CPn0081
                                          rpoC-RNA Polymerase Beta' -(CT314)
                         102221
  CPn0082
             98043
                                          tal-Transaldolase-(CT313)
             102332
                                    F
  CPn0083
                         103312
                                          predicted ferredoxin-(CT312)
                         103751
             103362
  CPn0084
                                          CT311 hypothetical protein
  CPn0085
             104506
                         103766
                                    R
                                          atpE-ATP Synthase Subunit E-(CT310)
                         105527
                                    F
             104904
  CPn0086
                                          CT309 hypothetical protein
                         106376
                                    F
  CPn0087
             105579
                                          atpA-ATP Synthase Subunit A-(CT308)
                         108145
             106373
  CPn0088
                                          atpB-ATP Synthase Subunit B-(CT307)
                                    F
                         109466
             108153
  CPn0089
                                          atpD-ATP Synthase Subunit D-(CT306)
                         110080
                                     F
  CPn0090
             109454
                                          atpI-ATP Synthase Subunit I-(CT305)
                         112053
                                    F
  CP50091
              110074
                                          atpK-ATP Synthase Subunit K-(CT304)
                         112573
                                     F
  CPN0092
              112151
                                          CT303 hypothetical protein
                         113015
                                     F
             112509
  CPh0093
                                          vals-Valyl tRNA Synthetase-(CT302)
                         115971
                                     F
              113152
   CP50094
                                          pknD-S/T Protein Kinase-(CT301)
   CP0095
                         118790
                                     F
              116037
                                          uvrA-Excinuclease ABC Subunit A-(CT333)
                         118837
                                     R
              124314
   Cpn0096
                                          pyk-Pyruvate Kinase-(CT332)
   CP10097
              124555
                         126006
                                     F
                                          htrB-Acyltransferase-(CT010)
                         126091
                                     R
   CPh0098
              127491
                         127865
                                     F
   Cgn0099
              127593
                                          CT011 hypothetical protein
                         127882
                                     R
   CPn0100
              129141
                                          ybbP family hypothetical protein-(CT012)
                         129141
                                     R
   CPn0101
              129932
                                          cydA-Cytochrome Oxidase Subunit I-(CT013)
                         131466
                                     F
   CP50102
              130123
                                          cydB-Cytochrome Oxidase Subunit II-(CT014)
                         132511
                                     F
              131480
   C₽#0103
                                          CT017 hypothetical protein
                         132676
                                     R
   CPp0104
              133875
                                          CT016 hypothetical protein
                         134029
                                     R
              134847
   CPn:0105
                                          phoH-ATPase-(CT015)
                         136374
                                     F
   CP00106
              135091
                                          CT058 hypothetical protein_1
                         136392
                                     R
              137162
   CP50107
                         137303
                                     R
   CPn0108
              137857
                                           ileS-Isoleucyl-tRNA Synthetase-(CT019)
              138655
                          141783
                                     F
   CPn0109
                                           lepB-Signal Peptidase I-(CT020)
                                     R
                         141827
   CPn0110
              143734
                                           CT021 hypothetical protein
              144686
                          143934
                                     R
   CPn0111
                                           rl31-L31 Ribosomal Protein-(CT022)
                         145093
                                     F
   CPn0112
              144767
                                           pfrA-Peptide Chain Releasing Factor (RF-1)-(CT023)
   CPn0113
              145335
                          146405
                                     F
                                           hemK-A/G specific methylase-(CT024)
                                     F
                          147261
   CPn0114
              146398
                                           ffh-Signal Recognition Particle GTPase-(CT025)
                          148622
                                     F
   CPn0115
              147279
                                           rs16-S16 Ribosomal Protein-(CT026)
                          148972
                                     F
   CPn0116
              148616
                                           trmD-tRNA (guanine N-1)-Methyltransferase-(CT027)
                          150071
   CPn0117
              148989
                                           rl19-L19 Ribosomal Protein-(CT028)
                          150464
                                     F
              150102
   CPn0118
                                           rnhB_1-Ribonuclease HII_1-(CT029)
   CPn0119
              150523
                          151164
                                     F
                                           gmk-GMP Kinase-(CT030)
                          151778
   CPn0120
              151164
                                           CT031 hypothetical protein
                          152068
   CPn0121
              151778
                                           metG-Methionyl-tRNA Synthetase-(CT032)
              152071
                          153723
                                     F
   CPn0122
                                           recD_1-Exodeoxyribonuclease V (Alpha Subunit)_1-(CT033)
                          153774
----CPn0123-
              _155969_
                          158068
   CPn0124
              156614
                                     F
   CPn0125
              158096
                          158605
                                     F
                          161085
              158809
                                     F
   CPn0126
                                           ytfF-Cationic Amino Acid Transporter-(CT034)
   CPn0127
              162143
                          161130
                                     R
                                           bpl1-Biotin Protein Ligase-(CT035)
   CPn0128
              162277
                          163053
                                     F
                                           similarity to CT036
   CPn0129
              163717
                          163064
                                     R
   CPn0130
               164245
                          163751
                                     R
                          165580
                                     F
   CPn0131
               164549
   CPn0132
               165587
                          166561
                                     F
                                           CHLPS hypothetical protein-(CT109)
                          166564
                                     R
   CPn0133
               167334
                                           groEL_1-HSP-60_1-(CT110)
   CPn0134
               169098
                          167467
                                     R
                                           groES-10KDa Chaperonin-(CT111)
                          169143
                                     R
   CPn0135
               169448
                                           pepF-Oligopeptidase-(CT112)
   CPn0136
               171401
                          169569
                                     R
                                           ybgI-ACR family-(CT108)
                          171502
                                      R
   CPn0137
               172254
                                           hemL-Glutamate-1-semialdehyde-2,1-aminomutase-(CT210)
                          172700
   CPn0138
               174019
```

```
- CPn0139
            174656
                        174093
                                    R
                                         yqgE-(CT210)
                                    R
                                         yqdE-(CT212)
                        174673
            175110
 CPn0140
                                         rpiA-Ribose-5-P Isomerase A-(CT213)
                        175110
                                    R
 CPn0141
            175802
                        175816
                                    R
            176091
 CPn0142
                                         *yxjG_Bs_1 Hypothetical Protein
                        176214
                                    R
 CPn0143
            177335
                                         clpB-Clp Protease ATPase-(CT113)
                        180560
                                    F
            177963
 CPn0144
                                         CT114 hypothetical protein
                        182369
            180777
 CPn0145
            182613
                        183095
 CPn0146
            183225
                        183671
 CPn0147
                                         pknl-S/T Protein Kinase-(CT145)
                        185702
                                    F
            183846
 CPn0148
                        187700
                                         dnlJ-DNA Ligase-(CT146)
            185715
 CPn0149
                                         CT147 hypothetical protein
                                    F
            187834
                        192444
 CPn0150
                                         mhpA-Monooxygenase-(CT148)
                        192625
            194142
 CPn0151
                                         CT149 hypothetical protein
            195265
                        194318
                                    R
 CPn0152
                                         leuS-Leucyl tRNA Synthetase-(CT209)
            195433
                        197892
                                    F
 CPn0153
                                         gseA-KDO Transferase-(CT208)
 CPn0154
            197892
                        199202
                                    F
            199691
                        199488
 CPn0155
 CPn0156
            200117
                        199770
                                    R
            200723
                        200298
                                    R
 CPn0157
 CPn0158
            201430
                        200894
                                    R
            201772
                        201467
 CPn0159
                                    R
                                         pfkA_1-Fructose-6-P Phosphotransferase_1-(CT207)
 CPn0160
            203791
                        202127
                                    R
                        203798
                                    R
                                         predicted acyltransferase family-(CT206)
            204622
 CPn:0461
 CPn0162
            205828
                        204803
                                    R
 CPn 0163
                        206394
            206026
            206498
                        206998
 CP110164
                                   P
            206998
                        207582
 CPd0165
 СРп0166
            207630
                        207962
                        207977
 CPn0167
            208306
                                    R
 CP40168
            208641
                        208417
 CP:0169
            209501
                        208710
                                    R
 CP10170
            211026
                        210025
                                    R
                                          *guaA-GMP Synthase
 CPn0171
            212435
                        211149
                                    R
                                         *guaB/impD-Inosine 5'-monophosphase dehydrogenase (COOH-terminal region
            213177
                        212440
 CPn0172
 CPn 0173
            213987
                        213715
                                    R
 CPn 9174
            214257
                        214724
                                   F
 CPn<u>0</u>175
            214898
                        215275
                                         CT153 hypothetical protein
 CPn 0176
            215286
                        216518
                                    F
 CPri0177
            217459
                        216608
                                    R
 CPn 0178
            218052
                        217789
                                    R
 CPn0179
            218403
                        218056
                                    R
 CPn0180
            218851
                        218355
                                   R
 CPn0181
            219175
                        218777
                                   R
                                         accC-Biotin Carboxylase-(CT124)
 CPn0182
            220695
                        219334
                                   R
 CPn0183
            221195
                        220695
                                         accB-Biotin Carboxyl Carrier Protein-(CT123)
                                   R
                                         efp_1-Elongation Factor P_1-(CT122)
 CPn0184
            221775
                        221221
                                    R
 CPn0185
            222451
                        221765
                                   R
                                         rpe/araD-Ribulose-P Epimerase-(CT121)
                                         *similarity to Cps IncA_1-(CT119)
 CPn0186
            222899
                        224068
                                   F
 CPn0187
            224248
                        225045
                                   F
                                         predicted methylase-(CT133)
 CPn0188
            225111
                        226400
                                         CT132 hypothetical protein
                                         CT131 homolog-(Possible Transmembrane Protein)
 CPn0189
            226400
                        229825
                                   F
 CPn0190
            229919
                        231274
                                         glnQ-ABC Amino Acid Transporter ATPase-(CT130)
 CPn0191
            231991
                        231314
                                   R
                                         glnP-ABC Amino Acid Transporter Permease-(CT129)
 CPn0192
                        231984
            232634
                                         *argR-Arginine Repressor
 CPn0193
            233126
                        232686
                                    R
                                        gcp_2-O-Sialoglycoprotein Endopeptidase_2-(CT197)
 CPn0194_
                        234241
            233210...
                                         oppA_1-Oligopeptide Binding Protein_1
 CPn0195
            234190
                        235785
                                         oppA_2-Oligopeptide Binding Protein_2-(CT198)
 CPn0196
            235939
                        237519
                                         oppA_3-Oligopeptide Binding Protein_3
 CPn0197
                        238882
            237578
                                   F
                                         oppA_4-Oligopeptide Binding Protein_4
. CPn0198
            239169
                        240746
                                         oppB_1-Oligopeptide Permease_1-(CT199)
 CPn0199
            241042
                        241983
                                   F
                                         oppC_1-Oligopeptide Permease_1-(CT200)
 CPn0200
            242017
                        242868
                                         oppD-Oligopeptide Transport ATPase-(CT201)
 CPn0201
            242864
                        243715
 CPn0202
                        244500
                                    F
                                         oppF-Oligopeptide Transport ATPase-(CT202)
            243715
 CPn0203
            245008
                        245802
 CPn0204
                        246002
            245817
 CPn0205
            246133
                        246327
 CPn0206
                        247161
                                    F
                                         CT203 hypothetical protein
            246409
                                         ybhI/sodiT1-Oxoglutarate/Malate Translocator-(CT204)
 CPn0207
            247208
                        248617
                                         pfkA_2-Fructose-6-P Phosphotransferase_2-(CT205)
 CPn0208
                        250602
            248953
 CPn0209
            251036
                        251272
```

```
CPn0210
            252384
                        251440
                                    R
 CPn0211
            252756
                        252463
                                    R
            254066
                        252888
 CPn0212
                                    R
 CPn0213
            254342
                        254190
                                    R
                        254446
 CPn0214
            255657
                                    R
 CPn0215
            257015
                        255759
 CPn0216
            257608
                        257174
                                   R
            257896
                        258579
                                   F
                                         ypdP-(CT140)
 CPn0217
 CPn0218
            259058
                        258582
                                   R
            259357
                        260472
                                         tgt-Queuine tRNA Ribosyl Transferase-(CT193)
                                   F
 CPn0219
 CPn0220
            260696
                        261238
                                   F
            261657
                        262064
                                   F
 CPn0221
            262504
                        262842
                                   F
                                         *weak similarity to Bacteriophage CHP1 (Orf4)
 CPn0222
                        263333
            262956
 CPn0223
                                   F
 CPn0224
            263435
                        263674
                                   F
            263873
                        264541
                                   F
 CPn0225
 CPn0226
            264566
                        264967
                                   F
            265416
                        265009
                                         dsbB-Disulfide bond Oxidoreductase-(CT176)
 CPn0227
                                   R
 CPn0228
            266110
                        265412
                                   R
                                         dsbG-Disulfide Bond Chaperone-(CT177)
                                         CT178 hypothetical protein
            266328
                        267560
 CPn0229
                                   F
 CP110230
            268253
                        267576
                                   R
                                         CT179 hypothetical protein
                                         tauB-ABC Transport ATPase (Nitrate/Fe)-(CT180)
 CP50231
            268957
                        268253
                                   R
                                         *similarity to 5'-Methylthioadenosine / S-Adenosylhomocysteine
 CPm0232
            270122
                        269232
                                   R
                                            Nucleosidase
 CPn0233
            270424
                        270248
                                   R
 CPm0234
            271240
                       270548
                                   R
                                         CT181 hypothetical protein
 CPn0235
            271416
                       272177
                                         kdsB-deoxyoctulonosic Acid Synthetase-(CT182)
                                   F
 CPn0236
            272156
                       273766
                                         pyrG-CTP Synthetase-(CT183)
                                         yggF Family-(CT184)
CPH0237
            273762
                       274214
                                   F
 CPn0238
            274303
                       275838
                                   F
                                         zwf-Glucose-6-P Dehyrogenase-(CT185)
 CPn0239
            275899
                       276672
                                         devB-Glucose-6-P Dehyrogenase (DevB family)-(CT186)
                                   F
CPn0240
            277861
                       276698
                                   R
CPn0241
            279354
                       278203
                                   R
CP#0242
            279918
                       279487
                                   R
CPm0243
            280555
                       280133
                                   R
CPn0244
            280918
                       281556
                                         adk-Adenylate Kinase-(CT128)
CPn0245
            281645
                       282499
                                         ydhO-Polysaccharide Hydrolase-Invasin Repeat Family-(CT127)
                                   F
CPn0246
            282952
                       282551
                                   R
                                         rs9-S9 Ribosomal Protein-(CT126)
CP20247
            283415
                       282969
                                        rl13-L13 Ribosomal Protein-(CT125)
                                  R
CPn0248
            284327
                       283650
                                   R
                                        ycfV/ybbA-ABC Transporter ATPase-(CT152)
CPn0249
            285841
                       284333
                                  R
                                        CT151 hypothetical protein
CPn0250
            286057
                       285902
                                  R
                                        rl33-L33 Ribosomal Protein-(CT150)
CPn0251
            286060
                       287559
                                  F
                                         *conserved hypothetical protein
CPn0252
            288112
                       287576
                                  R
                                        CT144 hypothetical protein (frame-shift with 0253?)
CPn0253
                       287950
            288456
                                  R
                                        CT144 hypothetical protein_1
CPn0254
            289262
                                        CT143 hypothetical protein_1
                       288459
                                  R
CPn0255
           290165
                       289329
                                  R
                                        CT142 hypothetical protein_1
CPn0256
                                        CT144 hypothetical protein_2
           291264
                       290398
                                  R
CPn0257
           292127
                       291267
                                  R
                                        CT143 hypothetical protein_2
CPn0258
           292534
                                        CT142 hypothetical protein (frame-shift with 0259?)
                       292133
                                  R
CPn0259
           292986
                       292441
                                  R
                                        CT142 hypothetical protein_2
CPn0260
           294045
                       293548
                                  R
                                        secA_1-Protein Translocase Subunit_1-(CT141)
CPn0261
           294302
                       295033
                                        ydaO-PP-Loop Superfamily ATPase-(CT217)
                                  F
CPn0262
           295091
                       295933
                                  F
                                        surE-SurE-like Acid Phosphatase-(CT218)
CPn0263
           296249
                       297136
                                  F
                                        yqfU hypothetical protein-(CT221)
CPn0264
           297730
                       297155
                                  R
                                        ubiD-Phenylacrylate Decarboxylase-(CT220)
CPn0265
           298620
                       297730
                                        ubiA-Benzoate Octaphenyltransferase-(CT219)
CPn0266
           299184
                       299876
                                  F
CPn0267
           300122
                       300910
                                  F
CPn0268
           300935
                       301318
                                  F
CPn0269
           302450
                       301476
                                  R
                                        Dipeptidase-(CT138)
CPn0270
           303325
                       302468
                                        ywlC-SuA5 Superfamily-related Protein-(CT137)
                                  R
CPn0271
           303634
                       304362
                                  F
                                        Lysophospholipase esterase-(CT136)
CPn0272
           305233
                       304340
                                  R
                                        dnaX_2-DNA Pol III Gamma and Tau_2-(CT187)
CPn0273
           305844
                       305227
                                        tdk-Thymidylate Kinase-(CT188)
                                  R
CPn0274
           308353
                       305852
                                  R
                                        gyrA_1-DNA Gyrase Subunit A_1-(CT189)
CPn0275
           310786
                       308372
                                  R
                                        gyrB_1-DNA Gyrase Subunit B_1-(CT190)
CPn0276
           311137
                       310793
                                  R
                                        CT191 hypothetical protein
CPn0277
           311910
                       311404
                                  R
CPn0278
           312875
                       312060
                                        *conserved outer membrane lipoprotein protein
CPn0279
           313537
                       312875
                                  R
                                        *Possible ABC Transporter Permease Protein
CPn0280
           314572
                       313550
                                        dppF-Dipeptide Transporter ATPase-(CT689)
```

```
dhnA-Predicted 1.6-Fructose Biphosphale Aldolase (dehydrin family) -
 CPn0281
            315057
                       316103
                                        xasA/gadC-Amino Acid Transporter-(CT216)
            316126
                       317529
                                  F
. CPn0282
            318497
                       317532
 CPn0283
            319045
                       318551
 CPn0284
                       319051
            320595
                                  R
 CPn0285
                                        mgtE-Mg++ Transporter (CBS Domain) - (CT194)
 CPn0286
            322059
                       320650
                                  R
                       322089
 CPn0287
            324221
                                        CT195 hypothetical protein
 CPn0288
            325716
                       324571
                                        aaaT-Neutral Amino Acid (Glutamate) Transporter-(CT230)
                       326996
                                  F
            325812
 CPn0289
                                        Na-dependent Transporter-(CT231)
            327042
                       328523
 CPn0290
                                        incB-Inclusion Membrane Protein B-(CT232)
            328667
                       329194
                                  F
 CPn0291
                                        incC-Inclusion Membrane Protein C-(CT233)
            329228
                       329836
 CPn0292
                                        CT234 hypothetical protein
            329949
                       332723
                                  F
 CPn0293
                                        cAMP-Dependent Protein Kinase Regulatory Subunit-(CT235)
            333092
                       333502
                                  F
 CPn0294
                                        acpP-Acyl Carrier Protein-(CT236)
            333863
                       333627
                                  R
 CPn0295
                                        fabG-Oxoacyl (Carrier Protein) Reductase-(CT237)
            334765
                       334022
 CPn0296
                                        fabD-Malonyl Acyl Carrier Transcyclase-(CT238)
            335697
                       334774
                                  :3
 CPn0297
                                        fabH-Oxoacyl Carrier Protein Synthase III-(CT239)
            336721
                       335717
                                  R
 CPn0298
                                        recR-Recombination Protein-(CT240)
 CPn0299
            336816
                       337415
                                  17
                                        yaeT-Omp85 Analog-(CT241)
            337783
                       340152
 CPn0300
                                        (OmpH-Like Outer Membrane Protein) - (CT242)
 CPn0301
            340250
                       340762
                                  1:
                                        lpxD-UDP Glucosamine N-Acyltransferase-(CT243)
            340787
                       341866
                                  ŀ.
 dPh0302
                                        CT244 hypothetical protein
 ⊊₹n0303
            342958
                       341921
                                  1
                                        pdhA/odpA-Pyruvate Dehydrogenase Alpha-(CT245)
            343133
                       344158
                                  F
 CPn0304
                                        pdhB/odpB-Pyruvate Dehydrogenase Beta-(CT246)
 ÇPn0305
            344154
                       345137
                                  1
                                        pdhC-Dihydrolipoamide Acetyltransferase-(CT247)
                       346431
 c Ph0306
            345145
                                  1
                                        glgP-Glycogen Phosphorylase-(CT248)
            348986
                       346515
                                  1:
 EPh0307
                                        similarity to CT249
            349234
                       349596
                                  F.
 @Pn0308
                                        dnaA_1-Replication Initiation Protein_1-(CT250)
 Ç₽n0309
            350974
                       349595
                                  R
                                        60IM-60kDa Inner Membrane Protein-(CT251)
 CPn0310
          353433
                       351049
                                  R
                                        lgt-Prolipoprotein Diacylglycerol Transferase-(CT252)
 dPn0311
            354438
                       353575
                                  R
            354524
                       354976
                                  F
                                        CT101 hypothetical protein
 @Pn0312
                                        acpS-Acyl-carrier Protein Synthase-(CT100)
 GPn0313
            354990
                       355355
                                        trxB-Thioredoxin Reductase-(CT099)
                       355353
                                  R
            356285
 @Pn0314
                                        rs1-S1 Ribosomal Protein-(CT098)
                       358716
  Pn0315
            356977
                                  F
                                        nusA-N Utilization Protein A-(CT097)
 ₫₱n0316
            358820
                       360121
                                  F
                                        infB-Initiation Factor-2-(CT096)
                       362750
 Ģ₽n0317
            360081
                       363126
                                        rbfA-Ribosome Binding Factor A-(CT095)
            362767
                                  F
 Ç₽n0318
                                        truB-tRNA Pseudouridine Synthase-(CT094)
 CPn0319
            363175
                       363879
                                  F
            363860
                                        ribF-FAD Synthase-(CT093)
 Ĉ₱ħ0320
                       364783
                                  F
                                       ychF-GTP Binding Protein-(CT092)
                       364767
 CPn0321
            365858
                                        yscu-YopS Translocation Protein U - (CT091)
 CPn0322
            366249
                       367328
                                  F
                                        lcrD- Low Calcium Response D-(CT090)
 CPn0323
            367331
                       369460
                                  F
                                        lcrE- Low Calcium Response E-(CT089)
                       370688
 CPn0324
            369492
                                  F
                                        sycE-Secretion Chaperone-(CT088)
 CPn0325
            370708
                       371148
                                       malQ-Glucanotransferase-(CT087)
 CPn0326
            371148
                       372725
                                  F
                                        r128-L28 Ribosomal Protein-(CT086)
 CPn0327
            372945
                       373211
                                  F
 CPn0328
            373241
                       374992
                                  F
                                        CT085 hypothetical protein
                                        Phopholipase D Superfamily [leader (33) peptide]-(CT084)
 CPn0329
            375088
                       376146
                                  F
                                        CT083 hypothetical protein
 CPn0330
            376675
                       376202
                                  R
                                        CT082 hypothetical protein
 CPn0331
            378437
                       376701
                                  R
                                        CHLTR T2 Protein-(CT081)
 CPn0332
            378655
                       378536
                                  R
 CPn0333
            379090
                       378800
                                        ltuB-(CT080)
                                  R
 CPn0334
            379311
                       379823
                                  F
                                        CT079 similarity
 CPn0335
                                        folD-Methylene Tetrahydrofolate Dehydrogenase-(CT078)
            379817
                       380674
                                  F
                                       _CPn0336 ___380650 __
                       381591
                                 - F
 CPn0337
            382027
                       381575
                                        smpB- Small Protein B-(CT076)
                                  R
                                        dnaN-DNA Pol III (beta chain)-(CT075)
CPn0338
            382278
                       383375
 CPn0339
            383420
                       384034
                                  F
                                        recF-ABC superfamily ATPase-(CT074)
                                        (frame-shift with 0339)
CPn0340
            383842
                       384156
                                  F
 CPn0341
            384160
                                        (frame-shift with 0340)
                       384495
                                  F
                                        predicted OMP [leader (19) peptide]-(CT073)
CPn0342
            384622
                       385062
CPn0343
            284999
                                        (frame-shift with 0342?)
                       385595
                                  F
                                       yaeL-Metalloprotease-(CT072)
CPn0344
            387420
                       385558
                                  R
CPn0345
                                       yaeM-(CT071)
            388572
                       387436
                                  R
CPn0346
                                        troD/ytgD-Integral Membrane Protein-(CT070)
            389675
                       388704
                                  R
                                        troC/ytgC-Integral Membrane Protein-(CT069)
CPn0347
            391021
                       389678
                                  R
CPn0348
                                        troB/ytgB-ABC transporter ATPase-(CT068)
            391803
                       391027
                                  R
                                        troA/ytgA-Solute Protein Binding Family-(CT067)
CPn0349
            392770
                       391790
                                  R
 CPn0350
                                        CT066 hypothetical protein
            393181
                       393684
                                        adt_1-ADP/ATP Translocase_1-(CT065)
CPn0351
            393888
                       395432
```

```
CPn0352
           395574
                       396830
                                   F
           396893
                       397135
                                   F
CPn0353
           397167
                       398507
                                  F
CPn0354
           399889
                       398591
                                  R
CPn0355
           400459
                       400109
CPn0356
           401317
                       400469
                                  R
CPn0357 .
CPn0358
           401751
                       401578
                                  R
                       403817
                                        lepA-GTPase-(CT064)
           402012
                                  F
CPn0359
                                        gnd-6-Phosphogluconate Dehydrogenase-(CT063)
CPn0360
           405358
                       403922
                                  R
                       405382
                                        tyrs-tyrosyl tRNA Synthetase-(CT062)
           406647
                                  R
CPn0361
                                        fliA/rpsD-Sigma-28/WhiG Family-(CT061)
CPn0362
           407825
                       407055
                                  R
                                        flhA-Flagellar Secretion Protein-(CT060)
                       407943
           409688
CPn0363
                                  R
                       410238
                                        fer4-Ferredoxin IV-(CT059)
CPn0364
           409966
           410528
                       411544
CPn0365
                                  F
CPn0366
           411976
                       412440
                                  F
                       413836
CPn0367
           413102
                                  F
           413790
                       414107
CPn0368
                                        CT058 hypothetical protein_2
                       415562
CPn0369
                                  F
           414351
                                        CT058 hypothetical protein_3
CPn0370
           415800
                       416912
                                  F
                       417503
CPn0371
           417147
                                  F
CPn0372
           417687
                       418001
                                  F
                                        gcpE-(CT057)
CPT0373
           418380
                       420218
                                  F
CPn0374
           420218
                       420961
                                  F
                                        CT056 hypothetical protein
CP30375
           421121
                       421615
                                  F
           421854
                       422294
                                  F
CPn0376
                                        sucB_1-Dihydrolipoamide Succinyltransferase_1-(CT055)
CPn0377
           423438
                       422347
                                        sucA-Oxoglutarate Dehydrogenase-(CT054)
CP770378
           426168
                       423445
                                  R
                                        CT053 hypothetical protein
CP70379
           426322
                       426765
                                        hemN_1-Coproporphyrinogen III Oxidase_1-(CT052)
           426758
                       427876
CP0380
CPn0381
           429809
                       428037
                                        CT326 similarity
                                        yabC/yraL-SAM-Dependent Methytransferase-(CT048)
           430749
                       430036
CPn0382
                                  R
CPn0383
           431693
                       430749
                                  R
                                        CT047 hypothetical protein
                                        hctB-Histone-like Protein 2-(CT046)
CP#0384
           432377
                       431862
                                  Ŕ
                                        pepA-Leucyl Aminopeptidase A-(CT045)
CPn:0385
           434018
                       432522
           434525
                                        ssb-SS DNA Binding Protein-(CT044)
CPn0386
                       434046
                                  R
CPn0387
           435196
                       434699
                                        CT043 hypothetical protein
                                  R
                                        glgX-Glycogen Hydrolase (debranching)-(CT042)
                       437320
CPn0388
           435329
                                  F
CP00389
           438134
                       437319
                                        CT041 hypothetical protein
                                  R`
C₽∰0390
           439144
                       438134
                                        ruvB-Holliday Junction Helicase-(CT040)
                                  R
CPn0391
           439692
                       439510
                                  R
CPn0392
           439814
                       440383
                                        dcd-dCTP Deaminase-(CT039)
                                  F
CPn0393
           440379
                       440723
                                        CT038 hypothetical protein
                                        tlyC_1-CBS Domain protein (Hemolysin Homolog)_1-(CT256)
           440736
                     441968
CPn0394
                                  F
CPn0395
           441964
                       443175
                                        CT257 hypothetical protein
                                  F
                                       yhfo-NifS-related protein-(CT258)
CPn0396
           444353
                       443241
                                  R
                                        PP2C phosphatase family-(CT259)
CPn0397
           445115
                       444381
CPn0398
           445533
                      445700
                                  F
CPn0399
                                       CT253 hypothetical protein
           445879
                       446523
                                  F
CPn0400
           446536
                      447306
                                  F
                                       CT254 hypothetical protein
CPn0401
           447884
                       447495
                                       CT255 hypothetical protein
                                  R
CPn0402
                                       mutY-Adenine Glycosylase-(CT107)
           448994
                      447888
                                  R
CPn0403
           449015
                      449710
                                  F
                                       yceC-predicted pseudouridine synthetase family-(CT106)
CPn0404
           450887
                      449871
                                  R
CPn0405
           451739
                       450966
                                        CT105 hypothetical protein
                                  R
                                        fabI-Enoyl-Acyl-Carrier Protein Reductase-(CT104)
CPn0406
           451969
                      452865
CPn0407
                                       HAD superfamily hydrolase/phosphatase-(CT103)
           453742
                      452858
                                  R
                                       CT102-hypothetical protein
CPn0408
           454105
                      454581
                                 -F
CPn0409
           454645
                      455127
                                       CT260 hypothetical protein
                                  F
CPn0410
                                        dnaQ_1-DNA Pol III Epsilon Chain_1-(CT261)
           455123
                      455833
CPn0411
           455833
                      456609
                                       CT262 hypothetical protein
                                  F
CPn0412
           456590
                      457246
                                        CT263 hypothetical protein
                                  F
CPn0413
           459203
                      457227
                                       msbA-Transport ATP Binding Protein-(CT264)
                                  R
                                        accA-AcCoA Carboxylase/Transferase Alpha-(CT265)
CPn0414
           460143
                      459172
CPn0415
           461498
                      460221
                                       CT266 hypothetical protein
                                  R
CPn0416
           461856
                      461557
                                       himD/ihfA-Integration Host Factor Alpha-(CT267)
                                  R
                                       amiA-N-Acetylmuramoyl Alanine Amidase-(CT268)
CPn0417
           463035
                      462244
                                  R
                                       murE-N-Acetylmuramoylalanylglutamyl DAP Ligase-(CT269)
CPn0418
           464401
                       462953
CPn0419
           466834
                      464876
                                       pbp3- transglycolase/transpeptidase-(CT270)
                                  R
CPn0420
           467108
                      466824
                                  R
                                       CT271 hypothetical protein
CPn0421
           467998
                                       yabC-PBP2B Family methyltransferase-(CT272)
                      467108
                                  R
CPn0422
           468242
                       468784
                                       CT273 hypothetical protein
CPn0423
           468791
                      469216
                                       CT274 hypothetical protein
```

```
dnaA_2-Replication Initiation Factor_2-(CT275)
                                   F
                        470961
 CPn0424
            469612
                                         CT276 hypothetical proteins
            470980
                        471564
 CPn0425
                                         CT277 similarity
                        471536
                                   R
            472111
 CPn0426
                                         nqr2-NADH (Ubiquinone) Dehydrogenase-(CT278)
                                   F
                        473715
            472207
 CPn0427
                                         nqr3-NADH (Ubiquinone) Oxidoreductase, Gamma-(CT279)
                                   F
                        474681
            473722
 CPn0428
                                         nqr4-NADH (Ubiquinone) Reductase 4-(CT280)
                        475319
            474681
 CPn0429
                                         ngr5-NADH (Ubiquinone) Reductase 5-(CT281)
                        476093
 CPn0430
            475326
                        476151
                                   R
            476483
 CPn0431
                        476514
                                    R
 CPn0432
            476816
                                         gcsH-Glycine Cleavage System H Protein-(CT282)
            477273
                        476929
                                   R
 CPn0433
                                         CT283 hypothetical protein
                        477276
                                   R
 CPn0434
            479462
                                         Phospholipase D superfamily [uncleavable leader peptide]-(CT284)
            480902
                        479475
                                   R
 CPn0435
                                         lplA-Lipoate Protein Ligase-Like Protein-(CT285)
                        480902
                                   R
             481618
 CPn0436
                                         clpC-ClpC Protease-(CT286)
__CPn0437
             481816
                        484350
                                   F
                                         ycbF-PP-loop superfamily ATPase-(CT287)
                        484334
             485416
 CPn0438
                                    F
             485553
                        486077
 CPn0439
                        486740
                                    F
 CPn0440
             486105
                                         CT007 hypothetical protein
                        487838
                                    F
             486891
 CPn0441
                                         CT006 hypothetical protein
                        488528
                                    F
             488013
 CPn0442
                                         CT005 hypothetical protein
                        489979
                                    F
             488729
 CPn0443
                                         pmp_6-Polymorphic Outer Membrane Protein G/I Family
                        494507
                                    F
             490287
 CPn0444
                                         pmp_7-Polymorphic Outer Membrane Protein G Family
                                    F
             494772
                        497579
 CPri0445
                                         pmp_8-Polymorphic Outer Membrane Protein G Family
                                    F
                        500415
 CP50446
             497626
                                         pmp_9-Polymorphic Outer Membrane Protein G/I Family
                                    F
             500568
                        503351
 CPn0447
                                          *yxjG_Bs_2 Hypothetical Protein
                        503698
                                    R
             504810
  CP10448
                                         pmp_10-PMP_10 (Frame-shift with 0451)
                                    R
             507231
                        505330
  CPn0449
                                         pmp_10-Polymorphic Outer Membrane Protein G Family
                        507180
                                    R
  CP#:0450
             508112
                                         pmp_11-Polymorphic Outer Membrane Protein G Family
                                    F
             508275
                        511058
  CPn0451
                                         pmp_12-Polymorphic Outer Membrane Protein A/I Family (truncated)
                                    F
             511319
                        512860
  CPn0452
                                         pmp_13 -Polymorphic Outer Membrane Protein G Family
                                    F
                        516152
             513234
  CPh0453
                                         pmp_14-Polymorphic Outer Membrane Protein H Family
                                    F
             516182
                        519115
  CPn0454
             520348
                        519458
                                    R
  CPn0455
             521532
                        520327
                                    R
  CPn0456
                        522120
                                    R
             523865
  CPn0457
                                    R
  CP10458
             526320
                        524236
             527005
                        526619
  CPn0459
  CPn0460
             527840
                        526992
                                    R
                        527844
                                    R
  CPn0461
             528638
                        529037
  CPn0462
             531052
                                    R
                        531191
  CPm0463
             532357
                        532366
                                    R
             532842
  CPn0464
                        532871
                                    R
  CPn0465
             533212
                                          pmp_15-Polymorphic Outer Membrane Protein E Family
                        536537
                                    F
             533724
  CPn0466
                                          pmp_16-Polymorphic Outer Membrane Protein E Family
             536633
                        539434
  CPn0467
                                          pmp_17-Polymorphic Outer Membrane Protein E Family
                        540432
                                    F
  CPn0468
             539632
                                          pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0469)
             540399
                         541460
                                    F
  CPn0469
                                          pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0470)
             541357
                         542532
                                    F
  CPn0470
                                          pmp_18-Polymorphic Outer Membrane Protein E/F Family
  CPn0471
             542564
                         545401
                         545581
                                    R
  CPn0472
             547905
             549593
                         548070
                                    R
  CPn0473
                                          CT365 hypothetical protein
  CPn0474
             551573
                         549807
                                    R
                                          glgB-Glucan Branching Enzyme-(CT866)
                         551685
                                    R
  CPn0475
             553844
                                          CT865 hypothetical protein
  CPn0476
             554844
                         553858
                                    R
                                          *yqeV_Bs Hypothetical Protein
  CPn0477
             556106
                         554844
                                    R
                                          hflx-GTP Binding Protein-(CT379)
  CPn0478
             557625
                         556210
                                    R
                                          phnP-Metal Dependent Hydrolase-(CT380)
  CPn0479
             558425
                         557616
                                    R
                                    R ____CT383_hypothetical_protein_____
---CPn0480----559303--
                       __558650-
                         559339
                                    R
  CPn0481
             560946
                                          artJ-Arginine Periplasmic Binding Protein-(CT381)
             561737
                         560961
                                    R
  CPn0482
                         564964
                                    F
             561836
  CPn0483
                                          aroG-Deoxyheptonate Aldolase-(CT382)
                         565824
                                    F
  CPn0484
             564970
                                          CT382.1 hypothetical protein
                                    F
  CPn0485
                         566229
             566038
                                          *hypothetical proline permease
  CPn0486
              567784
                         566405
                                    R
                                          CT384 hypothetical protein
                                    R
  CPn0487
             569740
                         568112
                                          hitA-HIT Family Hydrolase-(CT385)
  CPn0488
              570096
                         569767
                                    R
                                          CT386 hypothetical protein
                                    R
  CPn0489
                         570096
              570965
                                          CT387 hypothetical protein
  CPn0490
              571279
                         573333
                                    F
                                          CT389 hypothetical protein
                                    R
  CPn0491
              574352
                         573336
  CPn0492
              574652
                         574804
                                    R
  CPn0493
              575004
                         574855
  CPn0494
              575364
                         575146
                                     R
                                          aspC-Aspartate Aminotransferase-(CT390)
  CPn0495
              575603
                         576793
```

```
CT391 hypothetical protein
               576793
                         577812
                                     F
   CPn0496
                                          CT388 hypothetical protein
                          577820
                                     R
               578089
   CPn0497
                         578085
               579035
                                     R
   CPn0498
               580359
                          579205
                                     R
   CPn0499
                                          pros-Prolyl tRNA Synthetase-(CT393)
                                     F
               580659
                          582362
   CPn0500
                                          hrcA-HTH Transcriptional Repressor-(CT394)
                                     F
                          583650
               582457
   CPn0501
                                          grpE-HSP-70 Cofactor-(CT395)
                                     F
   CPn0502
               583650
                          584201
                                           dnaK-HSP-70-(CT396)
                          586213
    CPn0503
               584234
                                          vacB-ribonuclease family-(CT397)
                                     F
                          588514
               586487
   CPn0504
                                           *3-methyladenine DNA glycosylase
                          589106
                                     F
               588519
    CPn0505
                                          CT421 hypothetical protein
                                     F
                          589840
    CPn0506
               589172
                                          CT421.1 hypothetical protein
                          590122
               589961
    CPn0507
                                          CT421.2 hypothetical protein
                                     F
                          590300
    CPn0508
               590142
                                           (predicted Metalloenzyme) - (CT422)
                          590808
                                     F
               590335
    CPn0509
                                           tlyC_2-CBS Domains (Hemolysin homolog)_2-(CT423)
                          591973
               590813
    CPn0510
                                          rsbV_1-Sigma Regulatory Factor_1-(CT424)
               592141
                          592488
                                     F
    CPn0511
                                           CT425 hypothetical protein
                                     F
                          594412
    CPn0512
               592553
                                          Fe-S oxidoreductase_1-(CT426)
                                     F
               594647
                          595753
    CPn0513
                                           CT427 hypothetical protein
                          596520
                                     F
               595729
    CPn0514
                                           ubiE-Ubiquinone Methyltransferase-(CT428)
               596492
                          597181
    CPn0515
                                     R
                          597255
               598814
    CPn0516
                          598795
                                     R
               599631
    CPn0517
                                           CT429 hypothetical protein
               600803
                          599832
                                     R
    CPa0518
                                           dapF-Diaminopimelate Epimerase-(CT430)
               601674
                          600904 .
    CPn0519
                                           clpP-CLP Protease-(CT431)
                                     R
                          601646
               602218
    CPn0520
                                           glyA-Serine Hydroxymethyltransferase-(CT432)
               603797
                          602241
                                     R
    CPn0521
                                           CT433 hypothetical protein
               603987
                          604655
                                     F
    CPn0522
                          605052
    CPn0523
               604723
                          606179
                                     F
               605103
    CPn0524
                                           CT398 hypothetical protein
                          607283
                                     F
    င<sup>ူ</sup>့ဥက္0525
               606522
                                           yrbH-GutQ/KpsF Family Sugar-P Isomerase-(CT399)
                          607710
                                     R
    CPh0526
               608696
                                           sucB_2-Dihydrolipoamide Succinyltransferase_2-(CT400)
                          608726
               609904
    CPA0527
                                           gltT-Glutamate Symport-(CT401)
                          609921
                                     R
    CPn0528
               611162
                                           ycaH-ATPase-(CT402)
                          611165
                                     R
               612259
    CPn0529
                                           spoU_1-rRNA Methylase_1-(CT403)
    CPn0530
               613254
                          612460
                                     R
                                           SAM dependent methyltransferase-(CT404)
                          613245
    CPn0531
               614069
                                           ribC/risA-Riboflavin Synthase-(CT405)
    C. 10532
               614674
                          614075
                                           CT406 hypothetical protein
                          615385
                                     F
               614930
    CPn0533
                                           dksA-DnaK Suppressor-(CT407)
    CPn0534
               615413
                          615784
                                     F
                                           lspA-Lipoprotein Signal Peptidase-(CT408)
                          616296
               615793
    CPM0535
                                           dagA_1-D-Ala/Gly Permease_1-(CT409)
                                     F
               616345
                          617691
    CP-0536
                                           CT814.1 hypothetical protein
               617833
                          618189
    CPn0537
                                           CT814 hypothetical protein
               618212
                          618511
                                     F
    CPn0538
                                           pmp_19-polymorphic outer membrane protein A Family -(CT412)
               618705
                          621545
    CPn0539
                                           pmp_20-polymorphic outer membrane protein B Family-(CT413)
                                     F
    CPn0540
               621694
                          626862
                                           Solute binding protein (-yebL-Synechocystis Adhesin Homolog)-(CT415)
               627170
                          628003
                                     F
    CPn0541
                                           ABC Transporter ATPase-(CT416)
                                     F
               628003
                          628737
    CPn0542
                                           (Metal Transport Protein) - (CT417)
    CPn0543
               628725
                          629603
                                           yhbZ-GTP binding protein-(CT418)
                                     R
               630529
                          629525
    CPn0544
                                           rl27-L27 ribosomal protein-(CT419)
               630884
                          630633
                                     R
    CPn0545
                                           rl21-L21 Ribosomal Protein-(CT420)
               631229
                          630912
                                     R
    CPn0546
                                           ygbB family-(CT434)
               631661
                          632188
    CPn0547
                                           cysJ-Sulfite Reductase-(CT435)
                                     R
    CPn0548
               633231
                          632191
                                           rs10-S10 Ribosomal Protein-(CT436)
                         • 633255
    CPn0549
               633569
                                     R
                                           fusA-Elongation Factor G-(CT437)
                          633580
    CPn0550
               635661
                                     R
                                           rs7-S7 Ribosomal Protein-(CT438)
               636168
                           635698
                                      R
    CPn0551
                                     R -- rs12-S12-Ribosomal Protein-(CT439)
637747
                           636812
                                      R
    CPn0553
                                           CT440 hypothetical protein
                637854
                           638141
                                     F
    CPn0554
                                           tsp-Tail-Specific Protease-(CT441)
                           640241
                638298
    CPn0555
                                           crpA-15kDa Cysteine-Rich Protein-(CT442)
                640912
                           640325
                                      R
    CPn0556
                                           omcB-60kDa Cysteine-Rich Outer Membrane Complex Protein-(CT443)
                           641194
                                      R
                642861
    CPn0557
                                           omcA-9kDa-Cysteine-Rich Outer Membrane Complex Lipoprotein-(CT444)
     CPn0558
                643300
                           643031
                                      R
                                           CT441.1 hypothetical protein
                           643927
    CPn0559
                643742
                                           gltX-Glutamyl-tRNA Synthetase-(CT445)
     CPn0560
                645612
                           644098
                                      R
                                            euo-CHLPS Euo Protein-(CT446)
                           645871
                                      R
                646404
     CPn0561
                                            *CHLPS 43 kDa protein homolog_1
     CPn0562
                648036
                           646918
                                      R
                                            recJ-ssDNA Exonuclease-(CT447)
                           648293
                                      R
     CPn0563
                650056
                                            secD&secF-Protein Export Proteins SecD/SecF (fusion)-(CT448)
                           650145
     CPn0564
                654350
                                           CT449 hypothetical protein
                           654533
                                      R
     CPn0565
                655630
                                           yaeS family-(CT450)
                           656890
                                      F
     CPn0566
                656141
                                            cdsA-Phosphatidate Cytidylytransferase-(CT451)
                           657817
     CPn0567
                656894
```

```
cdsA-Phosphatidate Cytidylytransferase-(CT452)
                      658464
           657817
CPn0568
                                        plsC-Glycerol-3-P Acyltransferase-(CT453)
                                  F
                       659099
CPn0569
           658464
                                        args-Arginyl tRNA Transferase-(CT454)
                       660789
                                  F
           659107
CPn0570
                                        mura-UDP-N-Acetylglucosamine Transferase-(CT455)
                       660749
CPn0571
           662122
                                        CT456 hypothetical protein
                       664616
                                  F
           662352
CPn0572
                                        yebC family-(CT457)
                       664691
                                  R
CPn0573
           665404
           665945
                       665394
                                  R
CPn0574
                                        Yhhy-Amino Group Acetyl Transferase-(CT458)
           666494
                       665982
                                  R
                                        prfB-Peptide Chain Release Factor 2 (natural UGA frame-shift )-(CT459
CPn0575
                       666494
           667543
CPn0576
                                        prfB-(natural UGA frame-shift )
                                  R
                       667530
           667598
CPn0576
                                        SWIB (YM74) complex protein-(CT460)
           667895
                       668155
CPn0577
                                        yaeI-phosphohydrolase-(CT461)
                       669365
           668406
CPn0578
                                        ygbP/yacM-Sugar Nucleotide Phosphorylase-(CT462)
                       669993
           669361
CPn0579
                                        truA-Pseudouridylate Synthase I-(CT463)
                                  F
           669993
                       670793
CPn0580
                                        Phosphoglycolate Phosphatase-(CT464)
           671434
                       670745
CPn0581
                                        CT465 hypothetical protein
                       672177
           671503
CPn0582
                                        CT466 hypothetical protein
                       672717
           672400
CPn0583
                                        atoS/ntrB-2-Component Sensor-(CT467)
                       673798
                                  F
           672707
CPn0584
                                        *similarity to Cps IncA_2
                       673865
           675817
CPn0585
                                        atoC/ntrC-2-Component Regulator-(CT468)
                       677183
                                  F.
           676026
CPn0586
                                        *yvyD_Bs conserved hypothetical protein
                       678124
                                  F
CPn0587
           677441
                                        CT469 hypothetical protein
                       678626
                                  F
           678084
CPn0588
                                        CT470 hypothetical protein
                       679395
                                  F
           678640
CPa0589
                                        CT471 hypothetical protein
CPn0590
           680112
                       679516
                                        yagE family-(CT472)
                       681020
           680373
CPn0591
                                        yidD family-(CT473)
            681153
                       681461
CP00592
                                        CT474 hypothetical protein
                       681391
            682476
CPh0593
                                        pheT-phenylalanyl tRNA Synthetase Beta-(CT475)
            682583
                       684958
                                  F
CPA0594
                                        CT476 hypothetical protein
                       685926
           684958
CPn0595
                                        ada-methyltransferase-(CT477)
CPn0596
            685939
                       686457
                                  F
                                        oppC_2-Oligopeptide Permease_2-(CT478)
                       686479
                                  R
            688215
 CPh0597
                                        oppB_2-Oligopeptide Permease_2-(CT479)
c#h0598
            689697
                       688219
                                  R
                                        oppA_5-oligopeptide Binding Lipoprotein_5-(CT480)
            691802
                       689682
GPn0599
                                  R
CPn0600
            692147
                       691827
                                        CT483 hypothetical protein
                       692736
                                   R
            693053
 CPn0601
                                        CT484 hypothetical protein
            694105
                       693104
                                  R
 ල්ව්බ0602
                                        hemZ-Ferrochetalase-(CT485)
                       695185
            694205
 CPn0603
                                        fliy-Glutamine Binding Protein-(CT486)
                                  R
            695945
                       695196
 CPn0604
                                        yhhF-Methylase -(CT487)
            696707
                       696150
                                   R
 C₽n0605
                                        CT488 hypothetical protein
            697444
                       696707
                                   R
 CPh0606
                                        glgC-Glucose-1-P Adenyltransferase-(CT489)
            698895
                       697573
                                        *pyrF-Uridine 5 -Monophosphate Synthase (Ump Synthase)-truncated?
 CPh0607
            699645.
                       699016
                                   R
 CPn0608
                                        CT490 hypothetical protein
            699705
                       699986
                                   F
 CPn0609
                                        rho-Transcription Termination Factor-(CT491)
            701420
                       700029
                                   R
 CPn0610
                                        yacE-predicted phosphatase/kinase-(CT492)
                       701420
            702025
 CPn0611
                                        polA-DNA Polymerase I-(CT493)
            704631
                       702022
                                   R
 CPn0612
                                         sohB-Protease-(CT494)
                       704658
                                   R
            705656
 CPn0613
                                        adt_2-ADP/ATP Translocase_2-(CT495)
            707402
                       705783
                                   R
 CPn0614
                                         pgsA_1-Glycerol-3-P Phosphatidyltransferase_1-(CT496)
                       707634
 CPn0615
            708137
                                         dnaB-Replicative DNA Helicase-(CT497)
            708791
                       710137
                                   F
 CPn0616
                                         gidA-FAD-dependent oxidoreductase-(CT498)
                        712316
                                   F
            710484
 CPn0617
                                         lplA-Lipoate-Protein Ligase A-(CT499)
 CPn0618
            712306
                        713010
                                   F
                                        ndk-Nucleoside-2-P Kinase-(CT500)
            713444
                        713013
 CPn0619
                                         ruvA-Holliday Junction Helicase-(CT501)
 CPn0620
            714139
                        713519
                                   R
                                         ruvC-Crossover Junction Endonuclease-(CT502)
            714647
                        714144
 CPn0621
                                         CT503 hypothetical protein
 CPn0622
            715752
                        714793
                                   R
                                        -CT504-hypothetical-protein....
            716993
                       716163
 CPn0623
                                         gapA-Glyceraldehyde-3-P Dehyrogenase-(CT505)
 CPn0624
            718015
                        717011
                                   R
                                         rl17-L17 Ribosomal Protein-(CT506)
            718485
                        718060
                                   R
 CPn0625
                                         rpoA-RNA Polymerase Alpha-(CT507)
 CPn0626
            719616
                        718495
                                   R
                                         rsl1-Sl1 Ribosomal Protein-(CT508)
            720038
                        719640
                                   R
 CPn0627
                                         rs13-S13 Ribosomal Protein-(CT509)
 CPn0628
            720428
                        720063
                                   R
                                         secY-Translocase-(CT510)
                        720487
                                   R
 CPn0629
            721857
                                         rl15-L15 Ribosomal Protein-(CT511)
                        721885
                                   R
 CPn0630
             22316
                                         rs5-S5 Ribosomal Protein-(CT512)
                        722312
 CPn0631
            722806
                                         rll8-Ll8 Ribosomal Protein-(CT513)
                        722827
 CPn0632
            723195
                                         rl6-L6 Ribosomal Protein-(CT514)
                        723209
                                   R
 CPn0633
            723757
                                         rs8-S8 Ribosomal Protein-(CT515)
                        723787
 CPn0634
            724185
                                   R
                                         rl5-L5 Ribosomal Protein-(CT516)
                        724206
 CPn0635
            724745
                                   R
                                         rl24-L24 Ribosomal Protein-(CT517)
 CPn0636
            725082
                        724750
                                         rl14-L14 Ribosomal Protein-(CT518)
 CPn0637
            725464
                        725099
                                   R
                                         rs17-S17 Ribosomal Protein-(CT519)
 CPn0638
             725747
                        725490
```

F

```
rl29-L29 Ribosomal Protein-(CT520)
                          725743
              725958
   CPn0639
                                           rl16-L16 Ribosomal Protein-(CT521)
   CPn0640
              726377
                          725964
                                           rs3-S3 Ribosomal Protein-(CT522)
                          726409
              727077
   CPn0641
                                           rl22-L22 Ribosomal Protein-(CT523)
                          727096
              727428
   CPn0642
                                           rs19-S19 Ribosomal Protein-(CT524)
                          727450
              727713
   CPn0643
                                           rl2-L2 Ribosomal Protein-(CT525)
                          727722
              728573
   CPn0644
                                           rl23-L23 Ribosomal Protein-(CT526)
                          728598
              728930
   CPn0645
                                           rl4-L4 Ribosomal Protein-(CT527)
                          728950
                                     R
              729621
    CPn0646
                                           rl3-L3 Ribosomal Protein-(CT528)
                          729657
                                     R
   CPn0647
              730331
                                           CT529 hypothetical protein
                          730605
                                     Ŕ
              731603
    CPn0648
                                           fmt-Methionyl tRNA Formyltransferase-(CT530)
                          731710
                                     R
              732672
   CPn0649
                                           lpxA-Acyl-Carrier UDP-GlcNAc -(CT531)
                          732665
                                     R
              733501
    CPn0650
                                           fabZ-Myristoyl-Acyl Carrier Dehydratase-(CT532)
                          733517
                                     R
               733975
    CPn0651
                                           lpxC-Myristoyl GlcNac Deacetylase-(CT533)
                          733990
                                     R
              734835
    CPn0652
                                           cutE-Apolipoprotein N-Acetyltransferase-(CT534)
                          734868
                                     R
              736490
   CPn0653
                                           vdlD/yciA-acyl-CoA Thioesterase-(CT535)
                          736503
                                     R
              736967
    CPn0654
                                           dnaQ_2-DNA Pol III Epsilon Chain_2-(CT536)
                          737101
                                     R
               737847
    CPn0655
                          738048
                                     F
              737872
    CPn0656
                                           yjeE (ATPase or Kinase)-(CT537)
                          738051
                                     R
               738473
    CPn0657
                                           CT538 hypothetical protein
                          738455
                                     R
               739168
   · CPn0658
                                           trxA-Thioredoxin-(CT539)
                          739838
                                     F
               739533
    CPn0659
                                           spoU_2-rRNA Methylase_2-(CT540)
                          739860
                                      R
               740327
    CP10660
                                           mip-FKBP-type peptidyl-prolyl cis-trans isomerase-(CT541)
                          740327
                                      R
    CP10661
               741100
                                           aspS-Aspartyl tRNA Synthetase-(CT542)
               742923
                          741172
                                     R
    CPn0662
                                           hisS-Histidyl tRNA Synthetase-(CT543)
                          742901
                                      R
               744190
    CPn0663
                          744557
               744757
    CPn 0664
                                           uhpC-Hexosphosphate Transport -(CT544)
                                      F
                          746365
    CP10665
               745001
                                           dnaE-DNA Pol III Alpha-(CT545)
                          750107
                                      F
               746388
    CPn0666
                                           predicted OMP (leader (17)-(CT546)
                          750177
    CPn0667
               751058
                                           CT547 hypothetical protein
               751209
                          752162
    CP#0668
                                           CT548 hypothetical protein
                          752775
                                      F
    CPn0669
               752179
                                           rsbW-sigma regulatory factor-histidine kinase-(CT549)
               752765
                          753196
                                      F
    CPn0670
                                           CT550 hypothetical protein
                          753205
               753630
                                      R
    CPn0671
                                           dacF(pbp5)-D-Ala-D-Ala Caroxypeptidase-(CT551)
               753741
                          755048
    CPn0672
                                           CT552 hypothetical protein
                                      F
                          755463
               755287
    CPn0673
                                           fmu-RNA Methyltransferase-(CT553)
               756668
                          755577
                                      R
    CPT0674
                                           CT696 hypothetical protein
    C₽±0675
               757919
                          756768
                                      R
                                           homologous to CT695
    CP:10676
               759217
                          758051
                          759256
               760401
                                      R
    CP50677
    CPn0678
               761320
                          760682
                                      R
                                           pgk-Phosphoglycerate Kinase-(CT693)
                           761725
               762930
                                      R
    CPn0679
                                           ygo4-Phosphate Permease-(CT692)
               764248
                           762971
    CPn0680
                                           CT691 hypothetical protein
               764929
                          764258
                                      R
    CPn0681
                                            dppD-ABC ATPase Dipeptide Transport-(CT690)
               764984
                           765955
                                      F
    CPn0682
                                           dppF-ABC ATPase Dipeptide Transport-(CT689)
               765948
                          766919
                                      F
    CPn0683
                                            spoJ/parB-Chromosome Partitioning Protein-(CT688)
               768038
                          767181
    CPn0684
               768068
                           768217
                                      F
    CPn0685
    CPn0686
               768361
                           768176
                                      R
                                           CT482 hypothetical protein
    CPn0687
               768564
                           769214
                                      F
                                           CT481 hypothetical protein
               769382
                           770137
    CPn0688
                                           yfhO_1-NifS-related Aminotransferase_1-(CT687)
    CPn0689
               771404
                           770187
                                      R
                                            ABC Transporter Membrane Protein-(CT686)
               772680
                           771436
                                      R
    CPn0690
                                            abcX-ABC Transporter ATPase-(CT685)
    CPn0691
               773452
                           772685
                                      R
                                            ABC Transporter-(CT684)
               774912
                           773461
                                      R
    CPn0692
                                           TPR Repeats (O-Linked GlcNAc Transferase similarity)-(CT683)
    CPn0693
               776256
                           775240
                                      R
                                            pbp2-PBP2-transglycolase/transpeptidase-(CT682)
    CPn0694
               779599
                           776330
                                      R
                                           ompA-Major-Outer-Membrane-Protein-(CT681)
CPn0695
               780216
                          781382
                                      F
                                            rs2-S2 Ribosomal Protein-(CT680)
               781769
                           782599
    CPn0696
                                            tsf-Elongation Factor TS-(CT679)
    CPn0697
               782602
                           783447
                                      F
                                            pyrH-UMP Kinase-(CT679)
               783458
                           784201
                                      F
    CPn0698
                                            rrf-Ribosome Releasing Factor-(CT677)
    CPn0699
               784182
                           784721
                                      F
                                            CT676 hypothetical protein
                           785609
    CPn0700
               785097
                                      F
                                            karG-Arginine Kinase-(CT675)
    CPn0701
                           786672
                                      F
               785599
                                            yscC/gspD-Yop C/Gen Secretion Protein D-(CT674)
                           786929
                                      R
    CPn0702
               789685
                                            pkn5-S/T Protein Kinase-(CT673)
    CPn0703
               791190
                           789685
                                      R
                                            flin- Flagellar Motor Switch Domain/YscQ family-(CT672)
                           791209
    CPn0704
               792321
                                      R
                                            CT671 hypothetical protein
    CPn0705
               793173
                           792334
                                            CT670 hypothetical protein
    CPn0706
               793683
                           793180
                                      R
                                            yscN-Yop N (Flagellar-Type ATPase) - (CT669)
    CPn0707
               795029
                           793704
                                      R
                                            CT668 hypothetical protein
    CPn0708
                795705
                           795034
                                      R
                                            CT667 hypothetical protein
    CPn0709
               796188
                           795742
                                            CT666 hypothetical protein
    CPn0710
               796461
                           796210
```

```
CT665 hypothetical protein
 CPn0711
            796731
                       796486
                                   R
                                        FHA domain: homology to adenylate cyclase)-(CT664)
                       796781
                                   R
            799315
 CPn0712
                                        CT663 hypothetical protein
                       799332
            799721
                                   R
 CPn0713
                                        hemA-Glutamy1 tRNA Reductase-(CT662)
                       800091
                                   R
 CPn0714
            801107
                                        gyrB_2-DNA Gyrase Subunit B_2-(CT661)
                       803462
                                   F
            801657
 CPn0715
                                        gyrA_2-DNA Gyrase Subunit A_2-(CT660)
                       804902
                                   F
            803469
 CPn0716
                                        CT656 hypothetical protein
                       805306
 CPn0717
            805010
                                        CT657 hypothetical protein
                       805626
            805309
 CPn0718
                                        sfhB-(Pseudouridine Synthase)-(CT658)
                       806890
 CPn0719
            805916
                                        CT659 hypothetical protein
                       807236
            807003
 CPn0720
                                        kdsA-KDO Synthetase-(CT655)
                       808489
 CPn0721
            807683
                                        CT654 hypothetical protein
            808489
                       808974
 CPn0722
                                        yhbG-ABC Transporter ATPase-(CT653)
                       809703
                                   F
            808984
 CPn0723
            810527
                       809706
                                   R
CPn0724
                                        CT652.1 hypothetical protein
                       810587
                                   R
            810811
 CPn0725
                                        CT620 hypothetical protein
            813372
                       810880
                                   R
 CPn0726
                                        CT619 hypothetical protein
                       816192
                                   F
 CPn0727
            813577
                                        CHLPN 76kDa Homolog_1 (CT622)
            818477
                       816525
                                   R
 CPn0728
                                        CHLPN 76kDa Homolog_2 (CT623)
                       818592
            819857
                                   R
 CPn0729
                                        mviN-Integral Membrane Protein-(CT624)
 CPn0730
            821603
                       819963
                       821760
                                   F
 CPn0731
            821587
                                        nfo-Endonuclease IV-(CT625)
            822098
                       822976
                                   F
 CPn0732
                                        rs4-S4 Ribosomal Protein-(CT626)
 CP10733
            823727
                       823101
                                  R
                                        yceA-(CT627)
            823944
                       824915
 CPn-0734
                                         *pyrH/udk-Uridine Kinase (Uridine Monophosphokinase) (Pyrimidine
            825668
                       825003
                                   R
 CP10735
                                           Ribonucleoside Kinase).
                                        ygeD-Efflux Protein-(CT641)
                       825992
 CP10736
            827686
                                  R
                                        recC-Exodeoxyribonuclease V, Gamma-(CT640)
            827685
                       830756
                                   F
 CP10737
                                        recB-Exodeoxyribonuclease V, Beta-(CT639)
            830746
                       833895
                                   F
 CPn0738
                                        CT638 hypothetical protein
 CPp0739
                       833861
                                   R
            834871
                                        tyrB-Aromatic AA Aminotransferase-(CT637)
 CPH0740
            836048
                       834864
                                   R
                                        greA-Transcription Elongation Factor-(CT636)
 CPn0741
            838350
                       836185
                                   R
                                        CT635 hypothetical protein
 CPn0742
            838463
                       838888
                                   F
                                        ngrA-Ubiquinone Oxidoreductase, Alpha-(CT634)
 CPm0743
                       840362
                                   F
            838962
                                        hemB-Porphobilinogen Synthase-(CT633)
 CPn0744
            841384
                       840389
                                   R
            841903
                       841742
                                   R
 CPH0745
                                        CT632 hypothetical protein
 CPn0746
            841975
                       843567
                                   F
 CPn0747
            843675
                       843740
                                   F
                                        CT631 hypothetical protein
                                        CT631 hypothetical protein (frame-shift)
                                   F
CPn0747
            843725
                       843910
                                        ispA-Geranyl Transtransferase-(CT628)
 CPH0748
            844987
                       844121
                                        glmu-UDP-GlcNAc Pyrophosphorylase-(CT629)
 CP10749
            845629
                       845006
                                   R
                                        tctD/cpxR-HTH Transcriptional Regulatory Protein + Receiver Doman-
 CPn0750
            846411
                       845707
                                   R
                                           (CT630)
                                        CT651 hypothetical protein
 CPn0751
            846608
                       848434
                                   F
                                        recD_2-Exodeoxyribonuclease V, Alpha_2-(CT652)
 CPn0752
            848604
                       850082
                                   F
 CPn0753
            851006
                       850161
                                   R
                                        rs20-S20 Ribosomal Protein-(CT617)
 CPn0754
            851336
                       851040
                                   R
                                        CT616 hypothetical protein
 CPn0755
            851597
                       852799
                                        rpoD-RNA Polymerase Sigma-66 -(CT615)
 CPn0756
            852961
                       854676
                                   F
                                        folX-Dihydroneopterin Aldolase-(CT614)
 CPn0757
            854733
                       855134
                                        folP/dhpS-Dihydropteroate Synthase-(CT613)
 CPn0758
            855110
                       856459
                                   F
                                        folA-Dihydrofolate Reductase-(CT612)
 CPn0759
            856488
                       856997
                                   F
                                        CT611 hypothetical protein
 CPn0760
            856957
                       857694
                                   F
                                        CT610 hypothetical protein
 CPn0761
            857704
                       858375
                                   F
                                        recA-RecA recombination protein-(CT650)
 CPn0762
            859597
                       858539
                                   R
                                        ygfA-Formyltetrahydrofolate Cycloligase-(CT649)
 CPn0763
            860511
                       859972
                                   R
                                        CT648 hypothetical protein
-CPn0764 - 861807-
                       -860524
                                   R
                                        CT647 hypothetical protein
            862382
                       861801
 CPn0765
                                   R
                                        CT646 hypothetical protein
 CPn0766
            863782
                       862394
                                   R
                                        CT645 hypothetical protein
                       864177
 CPn0767
            863884
                                   F
                                        yohI/nir3-predicted oxidoreductase -(CT644)
 CPn0768
                       865163
                                   F
            864159
                                        topA-DNA Topoisomerase I-Fused to SWI Domain-(CT643)
 CPn0769
                       865121
                                   R
            867733
                                        CT642 hypothetical protein
 CPn0770
            868340
                       869131
                                   F
                                        rpoN-RNA Polymerase Sigma-54-(CT609)
                       869144
 CPn0771
            870463
                                   R
                                        uvrD-DNA Helicase-(CT608)
 CPn0772
            872385
                       870469
                                        ung-Uracil DNA Glycosylase-(CT607)
 CPn0773
                       873195
                                   F
            872488
 CPn0774
            873195
                       873425
                                   F
                                        CT606.1 hypothetical protein
 CPn0775
                                        yggV family-(CT606)
                       873414
            874031
                                   R
                                        CT605 hypothetical protein
 CPn0776
            874246
                       875487
                                        groEL_2-heat shock protein-60 -(CT604)
                       877178
                                   F
 CPn0777
            875601
                                         tsa/ahpC-Thio-specific Antioxidant (TSA) Peroxidase-(CT603)
 CPn0778
            877505
                       878092
                                   F
                                        CT602 hypothetical protein
 CPn0779
            878481
                       878095
```

```
papQ/amiB-N-Acetylmuramoyl-L-Ala Amidase-(CT601)
              879205
   CPn0780
                                           pal-Peptidoglycan-Associated Lipoprotein-(CT600)
                                     R
                          879198
   CPn0781
              879773
                                           tolB-polysaccharide transporter-(CT599)
                          879773
                                     R
              881065
   CPn0782
                                           CT598 hypothetical protein
                          881100
                                     R
   CPn0783
              881885
                                           exbD-Biopolymer Transport Protein-(CT597)
                          881892
                                     R
              882296
   CPn0784
                                           exbB/tolQ-polysaccharide transporter-(CT596)
                                     R
                          882296
   CPn0785
              882991
                                           dsbD/xprA-Thio:disulfide Interchange Protein-(CT595)
                          885293
                                     F
              883185
   CPn0786
                                           yabD/ycfH-PHP superfamily (urease/pyrimidinase) hydrolase-(CT594)
                                     F
              885619
                          886401
   CPn0787
                                           sdhC-Succinate Dehydrogenase-(CT593)
                          887432
                                     F
              886542
   CPn0788
                                           sdhA-Succinate Dehydrogenase-(CT592)
                                     F
                          889316
              887439
   CPn0789
                                           sdhB-Succinate Dehydrogenase-(CT591)
                          890103
                                     F
              889330
   CPn0790
                                           CT590 hypothetical protein
                                     R
                          890111
   CPn0791
              893050
                                           CT589 hypothetical protein
                          893108
   CPn0792
              894919
                                           rbsU-sigma regulatory family protein-PP2C phosphatase (RsbW
                          894919
                                     R
   CPn0793
              896823
                                             antagonist) - (CT588)
                                     F
                          898004
              897174
   CPn0794
                          899195
              898128
   CPn0795
                                     F
                          901340
   CPn0796
              899301
                          902694
                                     F
   CPn0797
              901600
                          903856
                                     F
              902846
   CPn0798
                          903940
   CPn0799
              904986
                                     R
                                           eno-Enolase-(CT587)
                          905249
                                     R
              906532
   CPn0800
                                           uvrB-Exinuclease ABC Subunit B-(CT586)
                          906727
                                     R
   CPq0801
              908697
                                           trpS-Tryptophanyl tRNA Synthetase-(CT585)
               909740
                          908709
                                     R
    CPn0802
                                           CT584 hypothetical protein
                          909752
                                     R
    CPn0803
              910303
                                           gp6D-CHLTR Plasmid Paralog-(CT583)
                          910310
                                     R
               911059
    Cirin 0804
                                           minD-chromosome partitioning ATPase-CHLTR plasmid protein GP5D-(CT582)
                          911067
                                     R
   C200805
              911831
                                           thrs-Threonyl tRNA Synthetase-(CT581)
                          911867
                                     R
               913771
    CPn0806
                                           CT580 hypothetical protein
                          914879
    CPn0807
               913971
                                           CT579 hypothetical protein
               916287
                          914956
                                     R
    CPh0808
                                           CT578 hypothetical protein
                          916307
                                     R
    C#10809
               917785
                                           CT577 hypothetical protein
               918184
                          917825
                                     R
    C-Pn0810
                                           lcrH_1-Low Ca Response Protein H_1-(CT576)
                          918208
                                     R
    CPn0811
               918900
                                           mutL-DNA Mismatch Repair-(CT575)
               919123
                          920862
                                     F
    CPn0812
                                           pepP-Aminopeptidase P-(CT574)
                          921934
                                     F
    CPn0813
               920870
                                           CT573 hypothetical protein
                          923357
                                      F
    CPn0814
               922107
                                           gspD/pilQ-Gen. Secretion Protein D-(CT572)
                          925622
                                      F
    CPn0815
               923361
                                           gspE-Gen. Secretion Protein E-(CT571)
                          927102
                                      F
    CPn0816
               925615
                                           gspF-Gen. Secretion Protein F-(CT570)
                          928287
                                     F
               927115
    CPn0817
                                           predicted OMP [leader (16) peptide]-(CT569)
                          928682
                                     F
    CPn0818
               928314
                                           CT568 hypothetical protein
                          929132
    CPn0819
               928689
                                           CT567 hypothetical protein
                          929659
    CPn0820
               929120
                                      F
                                           CT566 hypothetical protein
                        930668
                                      F
    CPn0821
               929667
                                           CT565 hypothetical protein
                          931229
                                      F
               930756
    CPn0822
                                           yscT/spaR-YopT Tranlocation T-(CT564)
                          931501
                                      R
    CPn0823
               932367
                                           yscs/fliQ-YopS/fliQ Translocation Protein-(CT563)
    CPn0824
               932662
                          932378
                                           yscR-Yop Translocation R-(CT562)
                          932677
                                      R
    CPn0825
               933594
                                           YSCL-Yop Translocation L-(CT561)
    CPn0826
               934310
                          933612
                                      R
                                           CT560 hypothetical protein
    CPn0827
               935264
                          934434
                                      R
                                           yscJ-Yop Translocation J-(CT559)
               936271
                          935267
    CPn0828
                                      F
    CPn0829
               936744
                          937298
               937444
                          937959
                                      F
    CPn0830
    CPn0831
               938267
                          938434
                                      F
                                           lipA-Lipoate Synthetase-(CT558)
    CPn0832
               939747
                          938827
                                           lpdA-Lipoamide Dehydrogenase-(CT557)
                          939747
    CPn0833
               941129
                                      R
                                           CT556 hypothetical protein
    CPn0834
               941553
                          942014
                                      F
                                           mot1=1-SWI/SNF-family=helicase=1-(CT555)
                                      -R---
CPn0835 945689
                         942045
                                            brnQ-Amino Acid (Branched) Transport-(CT554)
               946879
                          945722
    CPn0836
                                            nth-Enodnuclease III-(CT697)
    CPn0837
               947771
                          947145
                                      R
                                            thdF-Thiophene/Furan Oxidation Protein-(CT698)
               949106
                           947781
                                      R
    CPn0838
                                            psdD-Phosphatidylserine Decarboxylase-(CT699)
    CPn0839
               949257
                           950159
                                      F
                                            CT700 hypothetical protein
                           951544
    CPn0840
               950222
                                            secA_2-Translocase SecA_2-(CT701)
    CPn0841
                           954640
                                      F
               951731
                                            CT702 hypothetical protein (frame-shift with 0843)
                           954710
                                      R
    CPn0842
               954883
                                            CT702 hypothetical protein
                                      R
    CPn0843
               955191
                           954994
                                            yphC-GTPase/GTP-binding protein-(CT703)
                           955270
                                      R
    CPn0844
               956730
                                            pcnB_1-Poly A Polymerase_1-(CT704)
    CPn0845
               958079
                           956850
                                            clpX-CLP Protease ATPase-(CT705)
                           958112
                                      R
    CPn0846
               959374
                                            clpP-CLP Protease Subunit-(CT706)
    CPn0847
               959995
                           959387
                                      R
                                            tig/murI-Trigger Factor-peptidyl-prolyl isomerase-(CT707)
                           960177
    CPn0848
               961502
                                      R
                                            mot1_2-SWI/SNF family helicase_2-(CT708)
                           965285
    CPn0849
               961788
                                            mreB-Rod Shape Protein-Sugar Kinase-(CT709)
    CPn0850
               965293
                           966390
```

878591

R ·

```
CPn0851
             966396
                                         pckA-Phosphoenolpyruvate Carboxykinase-(CT710)
                         968195
                                         CT711 hypothetical protein
  CPn0852
             968316
                         970613
                                    F
                                         CT712 hypothetical protein
  CPn0853
             970637
                         971803
                                    F
                                         ompB-Outer Membrane Protein B-(CT713)
             972837
  CPn0854
                         971806
                                    Я
                                         gpdA-Glycerol-3-P Dehydrogenase-(CT714)
  CPn0855
             973995
                         972994
                                    R
                                         AgX-1 Homolog-UDP-Glucose Pyrophosphorylase-(CT715)
  CPn0856
             975377
                         973995
             975757
                         975392
                                         CT716 hypothetical protein
  CPn0857
                                    R
                                         fliI-Flagellum-specific ATP Synthase-(CT717)
             977055
                         975757
                                    R
  CPn0858
             977588
                         977055
                                         CT718 hypothetical protein
                                    R
  CPn0859
             978630
                        977608
                                         fliF-Flagellar M-Ring Protein-(CT719)
  CPn0860
                                         nifU-NifU-related protein-(CT720)
             979722
                        978925
                                    R
  CPn0861
             980873
                        979722
                                         yfhO_2-NifS-related protein_2-(CT721)
  CPn0862
                                    R
                                         pgmA-Phosphoglycerate Mutase-(CT722)
             981514
                        980831
  CPn0863
                                    R
             981670
                        982374
                                         yjbC-predicted pseudouridine synthase-(CT723)
  CPn0864
                                         CT724 hypothetical protein
             982418
                        982942
                                   F
  CPn0865
             983491
                        982916
                                         birA-Biotin Synthetase-(CT725)
  CPn0866
                                    R
                                         rodA-Rod Shape Protein-(CT726)
             983423
                        984667
                                   F
  CPn0867
                                         zntA/cadA-Metal Transport P-type ATPase-(CT727)
             986643
                        984670
  CPn0868
                                         CT728 hypothetical protein
             987401
                        986658
  CPn0869
                                   F.
  CPn0870
             988728
                        987448
                                   F.
                                         serS-Seryl tRNA Synthetase_2-(CT729)
                                         ribD-Riboflavin Deaminase-(CT730)
             988772
                        989899
  CPn0871
                                   F.
  CPn0872
             989963
                        991216
                                         ribA&ribB-GTP Cyclohydratase & DHBP Synthase -(CT731)
             991233
                                         ribE-Ribityllumazine Synthase-(CT732)
  @Pm0873
                        991694
                                   F
  CPn0874
             993107
                        991749
                                         CT733 hypothetical protein
  Ch0875
             993372
                                         CT734 hypothetical protein
                        994022
                                   F
                                         dagA_2-D-Alanine/Glycine Permease_2-(CT735)
  CPn0876
             994144
                        995517
  ₫₽£0877
             995533
                        995982
                                         ybcL family-(CT736)
  CPn0878
             996654
                        995992
                                   F
                                         SET Domain protein-(CT737)
  CPn0879
             997439
                        996645
                                         yycJ-metal dependent hydrolase-(CT738)
                                   R
  CPn0880
             999861
                        997444
                                         ftsK-Cell Division Protein FtsK-(CT739)
                                   R
  CPh0881
             1005667
                        1006209
  de 0882
             1006268
                        1007404
                                   F
             1008865
                        1007573
                                         dmpP/nqr6-Phenolhydrolase/NADH ubiquinone oxidoreductase-(CT740)
  CPn0883
                                   R
                                         CT741 hypothetical protein
  GPn0884
            1009359
                        1009009
                                   R
  CPn0885
            1010635
                                         ygcA-rRNA Methyltransferse-(CT742)
                        1009433
  CPh0886
            1011276
                                         hctA-Histone-Like Developmental Protein-(CT743)
                        1010908
                                   R
                                         CHLTR possible phosphoprotein-(CT744)
  C2n0887
             1011692
                        1014157
  CPn0888
                                         hemG-protoporphyrinogen Oxidase-(CT745)
            1015423
                        1014119
                                   R
                                         hemN_2-Coproporphyrinogen III Oxidase_2-(CT746)
  CE0889
            1016835
                        1015462
  CPn0890
                                         hemE-Uroporphyrinogen Decarboxylase-(CT747)
            1017805
                        1016819
                                   R
  CR.0891
             1021073
                        1017819
                                   R
                                         mfd-Transcription-Repair Coupling-(CT748)
 CPn0892
            1023661
                        1021046
                                   R
                                         alaS-Alanyl tRNA Synthetase-(CT749)
  CPn0893
            1023894
                        1025888
                                   F
                                         tktB-Transketolase-(CT750)
 CPn0894
            1026766
                        1025888
                                   R
                                         amn-AMP Nucleosidase-(CT751)
 CPn0895
             1026988
                        1027557
                                   F
                                         efp_2-Elongation Factor P_2-(CT752)
 CPn0896
            1027595
                        1027822
                                   F
                                         CT753 hypothetical protein
 CPn0897
            1028737
                        1027853
                                         (possible phosphohydrolase) - (CT754)
                                   R
 CPn0898
            1030460
                        1028904
                                   R
                                         Mitochondrial HSP60 Chaperonin Homolog-(CT755)
 CPn0899
            1030875
                        1032215
                                   F
                                        murF-Muramoyl-DAP Ligase-(CT756)
 CPn0900
            1032235
                        1033281
                                   F
                                         mraY-Muramoyl-Pentapeptide Transferase-(CT757)
 CPn0901
            1033287
                        1034537
                                   F
                                        murD-Muramoylalanine-Glutamate Ligase-(CT758)
 CPn0902
            1034543
                        1035241 ·
                                        nlpD-Muramidase (invasin repeat family) - (CT759)
                                   F
 CPn0903
            1035263
                                         ftsW-Cell Division Protein FtsW-(CT760)
                        1036417
                                   F
 CPn0904
            1036326
                        1037396
                                         murG-Peptidoglycan Transferase-(CT761)
                                   F
 CPn0905
            1037409
                                        murC&ddlA-Muramate-Ala Ligase & D-Ala-D-Alam Ligase-(CT762)
                        1039835
                                   F
 CPn0906
            1040340
                        1039915
                                        CT763 hypothetical protein
           _1040780____1040445____
CPn0907
                                   R ... *cutA Periplasmic Divalent Cation Tolerance Protein CutA (C-Type
                                           Cytochrome Biogenesis Protein)
 CPn0908
            1041589
                        1040780
                                   R
                                         CT764 hypothetical protein
                                        rsbV_2-Sigma Factor Regulator_2-(CT765)
 CPn0909
            1041637
                        1041966
                                   F
 CPn0910
            1041979
                                        miaA-tRNA Pyrophosphate Transferase-(CT766)
                        1043004
                                   F
 CPn0911
            1044043
                        1042985
                                   R
                                        Fe-S cluster oxidoreductase_2-(CT767)
 CPn0912
            1044129
                        1045760
                                   F
                                        CT768 hypothetical protein
 CPn0913
            1045760
                        1045945
                                   F
 CPn0914
            1045999
                        1046397
 CPn0915
                                        ybeB-iojap superfamily ortholog-(CT769)
            1046461
                        1046817
                                   F
                                         fabF-Acyl Carrier Protein Synthase-(CT770)
 CPn0916
            1046837
                        1048084
                                   F
 CPn0917
                                        hydrolase/phosphatase homolog-(CT771)
            1048090
                        1048539
                                   F
 CPn0918
            1049223
                                        ppa-Inorganic Pyrophosphatase-(CT772)
                        1048579
                                   R
 CPn0919
            1049378
                        1050430
                                         ldh-Leucine Dehydrogenase-(CT773)
                                   F
 CPn0920
                                         cysQ-Sulfite Synthesis/biphosphate phosphatase-(CT774)
            1051405
                        1050431
                                   R
 CPn0921
            1051535
                        1052293
                                         snGlycerol-3-P Acyltransferase-(CT775)
```

```
aas-Acylglycerophosphoethanolamine Acystransferase-(CT776)
            1052314
                       1053927
                                  F
 CPn0922
                                        bioF_1-Oxononanoate Synthase_1-(CT777)
 CPn0923
            1053984
                       1055093
                                  F
                                        priA-Primosomal Protein N' -(CT778)
 CPn0924
            1057274
                       1055028
                                        CT779 hypothetical protein
 CPn0925
            1057900
                       1057226
                                  R
            1058060
                       1058557
                                  F
                                        Thioredoxin Disulfide Isomerase-(CT780)
 CPn0926
            1059809
                       1058670
                                  R
                                        *CHLPS 43 kDa protein homolog_2
 CPn0927
            1061008
                       1059884
                                        *CHLPS 43 kDa protein homolog_3
                                  R
CPn0928
            1062292
                       1061186
                                  R
                                        *CHLPS 43 kDa protein homolog_4
CPn0929
                                  F
            1062857
                       1063330
CPn0930
CPn0931
            1064138
                       10657.18
                                  F
                                        lyss-Lysyl tRNA Synthetase-(CT781)
                                        cysS-Cysteinyl tRNA Synthetase-(CT782)
            1067142
                       1065721
CPn0932
                                  R
CPn0933
            1067535
                       1068578
                                        predicted disulfide bond isomerase-(CT783)
                                        rnpA-Ribonuclease P Protein Component-(CT784)
                       1068526
CPn0934
            1068942
                                  R
CPn0935
            1069091
                       1068957
                                        rl34-L34 Ribosomal Protein-(CT785)
                                        rl36-L36 Ribosomal Protein-(CT786)
            1069336
                       1069470
CPn0936
                                  F
CPn0937
           .1069496
                       1069798
                                        rs14-S14 Ribosomal Protein-(CT787)
                                        CT788 hypothetical protein -[leader (60) peptide-periplasmic]
                       1069849
           1070322
CPn0938
                                  R
CPn0939
            1070728
                       1071195
                                  F
                                        CT790 hypothetical protein
                                        uvrC-Excinuclease ABC, Subunit C-(CT791)
CPn0940
            1073012
                       1071204
                                  R
                                        mutS-DNA Mismatch Repair-(CT792)
CPn0941
           1075501
                       1073018
                                  R
                                        dnaG/priM-DNA Primase-(CT794)
           1075985
CPn0942
                       1077754
                                  F
CPn0943
           1077978
                       1078238
                                  F
                                        CT794.1 hypothetical protein
CP10944
           1078512
                       1078997
                                  F
CP;;0945
           1079070
                       1079660
                                        CT795 hypothetical protein
                                        glyQ-Glycyl tRNA Synthetase-(CT796)
CPn0946
           1082786
                       1079745
                                  R
                                        pgsA_2-Glycerol-3-P-Phosphatydyltransferase_2-(CT797)
CP10947
           1083442
                       1084059
                                  F
CPn0948
           1085474
                       1084047
                                        glgA-Glycogen Synthase-(CT798)
                                  R
CP10949
           1085929
                       1086483
                                        ctc-General Stress Protein-(CT799)
CPn0950
           1086488
                       1087027
                                        pth-Peptidyl tRNA Hydrolase-(CT800)
CPn0951
CPn0952
           1087122
                       1087457
                                        rs6-S6 Ribosomal Protein-(CT801)
                                  F
           1087478
                       1087723
                                        rs18-S18 Ribosomal Protein-(CT802)
CPn0953
           1087742
                       1088248
                                       rl9-L9 Ribosomal Protein-(CT803)
CPn0954
           1088286
                       1088708
                                       ychB-Predicted Kinase-(CT804)
CPn0955
           1088612
                       1089175
                                        (frame-shift with 0954)
CPn0956
           1089560
                       1090909
                                  F
                                        CT805 hypothetical protein
                                        ide/ptr-Insulinase family/Protease III-(CT806)
CPh0957
           1093788
                       1090963
                                  R
CPh0958
           1094785
                       1093793
                                       plsB-Glycerol-3-P Acyltransferase-(CT807)
CPh0959
           1096343
                       1094799
                                        cafE-Axial Filament Protein-(CT808)
                                  R
CPn0960
           1096764
                       1097102
                                  F
                                       CT809 hypothetical protein
CPh0961
           1097118
                                       rl32-L32 Ribosomal Protein-(CT810)
                       1097297
                                  F
CPN0962
           1097316
                       1098275
                                       plsX-FA/Phospholipid Synthesis Protein-(CT811)
                                       pmp_21-Polymorphic Outer Membrane Protein D Family-(CT812)
CPn0963
           1098398
                       1103224
                                  F
CPn0964
           1104758
                       1103301
                                  R
CPn0965
           1106736
                                       lpxB-Lipid A Disaccharide Synthase-(CT411)
                      1104925
                                  R
CPn0966
           1108037
                       1106748
                                  R
                                       pcnB_2-PolyA Polymerase_2-(CT410)
CPn0967
           1108512
                      1109885
                                  F
                                       mrsA/pgm-Phosphoglucomutase-(CT815)
CPn0968
           1109895
                       1111721
                                  F
                                       glmS-Glucosamine-Fructose-6-P Aminotransferase-(CT816)
CPn0969
           1111812
                                       0969-tyrP_1-Tyrosine Transport_1-(CT817) tyrP_1-Tyrosine Transport_1-
                      1112999
                                  F
                                          (CT817)
CPn0970
           1113461
                                       0970-tyrP_2-Tyrosine Transport_2-(CT818) tyrP_2-Tyrosine Transport_2-
                      1114648
                                  F
                                          (CT818)
CPn0971
           1114702
                      1115415 '
                                 F
                                       yccA-Transport Permease-(CT819)
CPn0972
           1116299
                      1115430
                                       ftsY-Cell Division Protein FtsY-(CT820)
                                  R
CPn0973
           1116370
                      1117527
                                  F
                                       sucC-Succinyl-CoA Synthetase, Beta-(CT821)
CPn0974
           1117544
                       1118422
                                  F
                                       sucD-Succinyl-CoA Synthetase, Alpha-(CT822)
CPn0975
           1119104
                      1119637
                                  F
CPn0976
           1120082 ___1121185
                                  F
CPn0977
           1121371
                      1122402
CPn0978
           1122665
                      1123693
                                  F
CPn0979
           1123980
                      1125443
                                  F
                                       htrA-DO Serine Protease-(CT823)
CPn0980
           1126982
                      1125504
                                       *similarity to Saccharomyces serevisiae hypothetical 52.9KD protein
                                  R
CPn0981
           1127031
                      1129952
                                       Zinc Metalloprotease (insulinase family) - (CT824)
CPn0982
           1131194
                      1129962
                                 R
                                       yigN family-(CT825)
CPn0983
           1132000
                      1131206
                                       pssA-Glycerol-Serine Phosphatidyltransferase-(CT826)
CPn0984
           1132379
                                       nrdA-Ribonucleoside Reductase, Large Chain-(CT827)
                      1135510
CPn0985
           1135534
                                       nrdB-Ribonucleoside Reductase, Small Chain-(CT828)
                      1136571
CPn0986
           1136724
                      1137395
                                       yggH-predicted rRNA Methylase-(CT829)
CPn0987
           1137516
                      1138115
                                 F
                                       ytgB-like predicted rRNA methylase-(CT830)
CPn0988
           1138986
                      1138075
                                 R
                                       murB-UDP-N-Acetylenolpyruvoylglucosamine Reductase-(CT831)
CPn0989
           1139495
                      1139016
                                       CT832 hypothetical protein
                                 R
CPn0990
           1139883
                      1140440
                                       infC-Initiation Factor 3-(CT833)
CPn0991
           1140421
                      1140612
                                       rl35-L35 Ribosomal Protein-(CT834)
```

```
rl20-L20 Ribosomal Protein-(CT835)
                      1140996
                                  F
CPn0992
           1140634
                                        pheS-Phenylalanyl tRNA Synthetase, Alpha-(CT836)
                      1142030
           1141014
                                  F
CPn0993
                                        CT837 hypothetical protein
CPn0994
           1142398
                      1144440
                                  F
                                        CT838 hypothetical protein
                      1144415
                                  R
           1145512
CPn0995
                                        CT839 hypothetical protein
                       1145519
                                  R
           1146589
CPn0996
                                        mesJ-PP-loop superfamily ATPase-(CT840)
                      1147664
CPn0997
           1146708
                                        ftsH-ATP-dependent zinc protease-(CT841)
                       1150584
                                  F
           1147855
CPn0998
                                        pnp-Polyribonucleotide Nucleotidyltransferase-(CT842)
                       1150766
                                  R
           1152847
CPn0999
                                        rs15-S15 Ribosomal Protein-(CT843)
           1153157
                       1152891
                                  R
CPn1000
                                        yfhC-cytosine deaminase-(CT844)
                      1153869
                                  F
CPn1001
           1153405
                                        CT845 hypothetical protein
                       1154089
                                  F
           1153862
CPn1002
                                        CT846 hypothetical protein
                      1154092
                                  R
CPn1003
           1154796
                                        CT847 hypothetical protein
                                  R
                       1154879
           1155397
CPn1004
                                        CT848 hypothetical protein
                      1155415
                                  R
CPn1005
           1155933
                                        CT849 hypothetical protein
                       1155990
                                  R
           1156472
CPn1006
                                        CT849.1 hypothetical protein
                                  F
                       1156907
CPn1007
           1156689
                                        CT850 hypothetical protein
                       1158223
                                  F
CPn1008
           1156928
                                        map-Methionine Aminopeptidase-(CT851)
                       1158186
           1159058
                                  R
CPn1009
                                        CT852 hypothetical protein
           1159672
                       1159067
CPn1010
                                        CT853 hypothetical protein
                       1159902
                                  R
           1160306
CPn1011
                                        yzeB-ABC transporter permease-(CT854)
           1162193
                       1160421
CPn1012
                                        fumC-Fumarate Hydratase-(CT855)
           1162245
                       1163624
CPn1013
                                        ychM-Sulfate Transporter-(CT856)
           1165426
                       1163732
CPn 1-014
                                        CT857 hypothetical protein (possible IM protein)
                       1166893
                                  F
           1165634
CPn1015
                                        CT858 hypothetical protein
           1167042
                       1168898
CPn1016
                                        lytB-Metalloprotease-(CT859)
           1169006
                       1169935
                                  F
CPn1917
                       1170629
           1169898
CPn1018
                                        CT860 hypothetical protein
CPn1019
           1172128
                       1170638
                                  R
                                        CT861 hypothetical protein
                       1172150
                                  R
CPn1020
           1173679
                                        lcrH_2-Low Calcium Response_2-(CT862)
                       1173698
                                  R
CPn1021
           1174213
                                        CT863 hypothetical protein
                       1174216
           1175673
CPn 1022
                       1176331
                                  F
CPn1023
           1176035
                                        xerD-Integrase/recombinase-(CT864)
                       1176334
                                  R
           1177236
CPn 1024
                                        pgi-Glucose-6-P Isomerase-(CT378)
                       1178879
                                  F
CPn 1025
           1177302
                                        ltuA-(CT377)
           1178997
                       1179137
CPn1026
                       1180755
                                  F
CPn1027
           1179175
                                        mdhC-Malate Dehyrogenase-(CT376)
                       1181999
                                  F
           1181016
CPn1028
CPn 1029
           1182008
                       1182844
                                  F
                                        predicted D-amino acid dehyrogenase-(CT375)
                       1182843
                                  R
CPn1030
           1183886
                                        arcD-Arginine/Ornithine Antiporter-(CT374)
CPnF031
           1185552
                       1184098
                                  R
                                        CT373 hypothetical protein
                       1185566
                                  R
CPn1032
           1186150
                                        CT372 hypothetical protein
CPn1033
           1187500
                       1186187
                                  R
                                        Predicted OMP_1 (CT371) [leader (18) peptide]
                       1187732
                                  R
CPn1034
           1188517
                                        AroE-Shikimate 5-Dehyrogenase-(CT370)
                       1188570
CPn1035
           1190000
                                        AroB-Dehyroquinate Synthase-(CT369)
                       1189984
                                  R
CPn1036
           1191135
                                        AroC-Chorismate Synthase-(CT368)
                       1191123
CPn1037
           1192199
                                  R
                       1192199
                                        aroL-Shikimate Kinase II-(CT367)
                                  R
CPn1038
           1192726
                                        aroA-Phosphoshikimate Vinyltransferase-(CT366)
                       1192665
CPn1039
            1193999
                       1194073
                                  R
CPn1040
           1194741
                                         *bioA-Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase
                       1194726
CPn1041
            1195994
                                  R
                                         *bioD-dethiobiotin synthetase
                       1195934
CPn1042
           1196590
                                  R
                                        bioF_2-Oxononanoate Synthase_2
CPn1043
            1197717
                       1196572
                                   R
                                         *bioB-Biotin Synthase
                       1197699
                                  R
CPn1044
            1198691
                                         conserved hypothetical bacterial membrane protein
CPn1045
            1199590
                       1198901
                                         *Tryptophan Hyroxylase
                       1199590
(CPn1046"
            1200675
                                   R
                                         dapB-Dihydrodipicolinate Reductase-(CT364)
 CPn1047
            1200552 -
                       1201343
                                  F
                                         asd-Aspartate Dehydrogenase-(CT363)
-CPn1048----
           -1201606-
                      -1202604
                                  F
                                         lysC-Aspartokinase III-(CT362)
CPn1049
            1202595
                       1203914
                                   F
                                         dapA-Dihydrodipicolinate Synthase-(CT361)
CPn1050
            1203926
                       1204798
                                  F
CPn1051
            1204962
                       1205270
· CPn1052
            1205417
                       1206169
                                  F
CPn1053
                       1206701
                                   F
            1206153
CPn1054
            1207034
                       1209466
                                   F
                       1210521
CPn1055
            1209694
CPn1056
            1210527
                       1211228
                                   F
                                         CT356 hypothetical protein
                       1213596
                                   F
CPn1057
            1211497
                                         CT355 hypothetical protein
 CPn1058
            1213748
                       1214836
                                   F
                                         kgsA-Dimethyladenosine Transferase-(CT354)
CPn1059
                       1215678
            1214848
                                         dxs/tkt-Transketolase-(CT331)
 CPn1060
            1217658
                       1215727
                                         CT330 hypothetical protein
                                   R
 CPn1061
            1217920
                       1217666
                                         xseA-Exodoxyribonuclease VII-(CT329)
 CPn1062
            1219820
                       1218159
                                         tpiS-Triosephosphate Isomerase-(CT328)
                       1220712
 CPn1063
            1219951
```

CPn1064	1220719	1220895	F	·
CPn1065	1221095	1220928	R	
CPn1066	1221135	1221488	F	
CPn1067	1221735	1222292	F	def-Polypeptide Deformylase-(CT353)
CPn1068	1223258	1222365	R	rnhB_2-Ribonuclease HII_2-(CT008)
CPn1069	1223513	1223941	F	yfgA-HTH Transcriptional Regulator-(CT009)
CPn1070	1225511	1224144	R	
CPn1071	1227324	1225885	R	
CPn1072	1227969	1228835	F	
CPn1073	1229011	1229832	F	Predicted OMP_2 -(CT371)

Table 2 (Supplemental Data) Functional Assignments of *C. pneumoniae* Coding Sequences. *C. trachomatis* genes are shown in parentheses.

	5	Amino Acid Biosynthesis								
		Aromatic	Family							
		1039	(CT366)	aroA	Phosphoshikimate Vinyltransferase					
		1036	(CT369)	aroB	Dehyroquinate Synthase					
		1037	(CT368)	aroC	Chorismate Synthase					
	10	1035	(CT370)	aroÉ	Shikimate 5-Dehyrogenase					
		0484	(CT382)	aroG	Deoxyheptonate Aldolase					
		1038	(CT367)	aroL	Shikimate Kinase II					
		0740	(CT637)	tyrB	Aromatic AA Aminotransferase					
		Aspartate	Family (lys	ine)						
	15	1048	(CT363)	asd	Asp2:tate Dehydrogenase					
		1050	(CT361)	dapA	Dihydrodipicolinate Synthase					
O		1047	(CT364)	dapB	Dihydrodipicolinate Reductase					
٠O		0519	(CT430)	dapF	Diaπ inopimelate Epimerase					
a Pro		1049	(CT362)	lysC	Aspa tokinase III					
	20	Serine Fa	ımily							
Ø		0433	(CT282)	gcsH	Glyc ne Cleavage System H Protein					
ļ.		0521	(CT432)	giyA	Serine Hydroxymethyltransferase	•				
II)		. Base & N	lucleotide M	etabolism						
LT:		0171		guaA	GMP Synthase	•				
ā	25	0172		guaB	Inosine 5'-Monophosphase Dehydrogenase					
ŀ÷		0608			Uridine 5'-Monophosphate Synthase					
ŀ÷		0735			Uridine Kinase					
ļ.		0244	(CT128)	adk	Adenylate Kinase					
	20	0894	(CT751)	amn	AMP Nucleosidase					
.E	30	0568	(CT452)	cmk	CMP Kinase					
		0392	(CT039)	dcd	dCTP Deaminase					
		0059	(CT292)	dut	dUTP Nucleotidohydrolase					
		0120	(CT030)	gmk	GMP Kinase					
	35	0619	(CT500)	ndk	Nucleoside-2-P Kinase					
	55	0984	(CT827)	nrdA	Ribonucleoside Reductase, Large Chain Ribonucleoside Reductase, Small Chain					
		0985 0236	(CT828) (CT183)	nrdB pyrG	CTP Synthetase					
		0698	(CT678)	ругН	UMP Kinase					
		0273	(CT188)	tdk	Thymidylate Kinase					
	40	0659	(CT539)	trxA	Thioredoxin					
		0314	(CT099)	trxB	Thioredoxin Reductase					
		1001	(CT844)	yfhC	Cytosine Deaminase					
			(3.5)	,						
					Biosynthesis of Cofactors					
	45_	Biotin, L	poate-&-Ub	iquinone	Harring states and the contraction of the contraction of the contraction of the contraction of the contraction					
		1041		bioA	Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase					
		1044		bioB	Biotin Synthase					
		1042		bioD	Dethiobiotin Synthetase					
		0923	(CT777)	bioF_1	Oxononanoate Synthase_I	•				
	50	1043	(CT777)	bioF_2	Oxononanoate Synthase_2					
		0866	(CT725)	birA	Biotin Synthetase					
		0748	(CT628)	ispA	Geranyl Transtransferase					
		0832	(CT558)	lipA	Lipoate Synthetase					

```
0264
                          (CT220)
                                      ubiD
                                                Phenylacrylate Decarboxylase
                   0515
                          (CT428)
                                      ubiE
                                                Ubiquinone Methyltransferase
                 Folic Acid
         5
                                                Dihydrofolate Reductase
                   0759
                          (CT612)
                                      folA
                                                Methylene Tetrahydrofolate Dehydrogenase
                                      folD
                   0335
                          (CT078)
                                                Dihydropteroate Synthase
                                      folP
                   0758
                          (CT613)
                   0757
                          (CT614)
                                      fol X
                                                Dihydroneopterin Aldolase
                                                Formyltetrahydrofolate Cycloligase
                   0763
                          (CT649)
                                      ygfA
       10
                Porphyrin
                                                Glutamyl tRNA Reductase
                   0714
                          (CT662)
                                      hemA
                                                Porphobilinogen Synthase
                   0744
                          (CT633)
                                      hemB
                                                Porphobilinogen Deaminase
                   0052
                          (CT299)
                                      hemC
                          (CT747)
                                      hemE
                                                Uroporphyrinogen Decarboxylase
       15
                   0888
                          (CT745)
                                      hemG
                                                protoporphyrinogen Oxidase
                                                Glutamate-1-Semialdehyde-2,1-Aminomutase
                   0138
                          (CT210)
                                      hemL
                                                Coproporphyrinogen III Oxidase_I
                   0380
                          (CT052)
                                      hemN 1
                   0889
                          (CT746)
                                      hemN_2
                                                Coproporphyrinogen III Oxidase_2
Ferrochetalase
                   0603
                          (CT485)
                                      hemZ
       20
                 Riboflavin
                   0872 (CT731)
                                      ribA&ribB GTP Cyclohydratase & DHBP Synthase
                          (CT405)
                                      ribC
                                                Riboflavin Synthase
                   0532
                                                Riboflavin Deaminase
                   0871
                          (CT730)
                                      ribD
                   0873
                          (CT732)
                                      ribE
                                                Ribityllumazine Synthase
       25
                   0320
                          (CT093)
                                      пbF
                                                FAD Synthase
M
                 Cell Envelope
14
                 Fatty Acid & Phospholipid Metabolism
14
                   0161 (CT206)
                                                (predicted acyltransferase family)
]=
       30
                                                Acylglycerophosphoethanolamine Acyltransferase
                   0922
                         (CT776)
                                      aas
1=
                                                AcCoA Carboxylase/Transferase Alpha
                   0414
                          (CT265)
                                      accA
ű
                   0183
                          (CT123)
                                      accB
                                                Biotin Carboxyl Carrier Protein
Ū
                   0182
                          (CT124)
                                      accC
                                                Biotin Carboxylase
                   0058
                          (CT293)
                                      accD
                                                AcCoA Carboxylase/Transferase Beta
       35
                   0295
                          (CT236)
                                      acpP
                                                Acyl Carrier Protein
                   0313
                                                Acyl-carrier Protein Synthase
                          (CT100)
                                      acpS
                                                Phosphatidate Cytidylytransferase
                   0567
                          (CT451)
                                      cdsA
                   0297
                          (CT238)
                                      fabD
                                                Malonyl Acyl Carrier Transcyclase
                                                Acyl Carrier Protein Synthase
                   0916
                          (CT770)
                                      fabF
       40
                   0296
                          (CT237)
                                      fabG
                                                Oxoacyl (Carrier Protein) Reductase
                                                Oxoacyl Carrier Protein Synthase III
                   0298
                          (CT239)
                                      fabH
                                                Enoyl-Acyl-Carrier Protein Reductase
                   0406
                          (CT104)
                                      fabl
                   0651
                          (CT532)
                                      fabZ
                                                Myristoyl-Acyl Carrier Dehydratase
                   0098
                          (CT010)
                                      htrB
                                                Acyltransferase
       45
                   0271
                          (CT136)
                                                Lysophospholipase Esterase
                   0615
                          (CT496)
                                                Glycerol-3-P Phosphatidyltransferase_1
                                      pgsA_1
                   0947
                          (CT797)
                                                Glycerol-3-P Phosphatydyltransferase_2
                                      pgsA_2
                          -(CT807)-
                                      :plsB=
                                                Glycerol=3=P=Acyltransferase=
                                                Glyceroi-3-P Acyltransferase
                   0569
                          (CT453)
                                      plsC
       50
                                                FA/Phospholipid Synthesis Protein
                   0962
                          (CT811)
                                      plsX
                          (CT699)
                   0839
                                      psdD
                                                Phosphatidylserine Decarboxylase
                                                Glycerol-Serine Phosphatidyltransferase
                   0983
                          (CT826)
                                      pssA
                                                snGlycerol-3-P Acyltransferase
                   0921
                          (CT775)
                                                Acyl-CoA Thioesterase
                   0654
                          (CT535)
                                      yciA
       55
                   0877
                          (CT736)
                                                CT736 Hypothetical Protein
                                      ybcL
                 LPS
```

(CT219)

ubiA

0265

Benzoate Octaphenvitransferase

```
KDO Synthetase
                                   kdsA
                 0721
                        (CT655)
                                            Deoxyoctulonosic Acid Synthetase
                       (CT182)
                                   kdsB
                 0235
                                            Acyl-Carrier UDP-GlcNAc O-Acyltransferase
                                   lpxA
                       (CT531)
                 0650
        5
                                             Lipid A Disaccharide Synthase
                                   lpxB
                       (CT411)
                 0965
                                            Myristoyl GlcNac Deacetylase
                                   lpxC
                 0652
                       (CT533)
                                             UDP Glucosamine N-Acyltransferase
                                   ipxD
                 0302
                       (CT243)
               Membrane Proteins, Lipoproteins & Porins
                                             60kDa Inner Membrane Protein
                 0310
                        (CT251)
                                   60IM
                                             15kDa Cysteine-Rich Protein
      10
                 0556
                        (CT442)
                                   сгрА
                                             Apolipoprotein N-Acetyltransferase
                                   cutE
                  0653
                        (CT534)
                                             Prolipoprotein Diacylglycerol Transferase
                  0311
                       (CT252)
                                   lgt
                                             9kDa-Cysteine-Rich Lipoprotein
                  0558
                        (CT444)
                                   omcA
                                             60kDa Cysteine-Rich OMP
                        (CT443)
                                   omcB
                                             Major Outer Membrane Protein
      15
                        (CT681)
                                   ompA
                  0695
                                             Outer Membrane Protein B
                        (CT713)
                                   ompB
                  0854
                                             Peptidoglycan-Associated Lipoprotein
                        (CT600)
                  0781
                                   pal
                                             Omp85 Homolog
                  0300
                        (CT241)
                                   yaeT
               Peptidoglycan
      20
                  0417 (CT268)
                                   amiA
                                             N-Acetylmuramoyl Alanine Amidase
N-Acetylmuramoyl-L-Ala Amidase
                  0780
                        (CT601)
                                   amiB
                                             D-Ala-D-Ala Caroxypeptidase
                  0672
                        (CT551)
                                   dacF
                                             Glucosamine-Fructose-6-P Aminotransferase
                        (CT816)
                                    glmS
                  0968
Ü
                                             UDP-GlcNAc Pyrophosphorylase
                  0749
                        (CT629)
                                    glmU
                                             Muramoyl-Pentapeptide Transferase
       25
l-h
                  0900
                        (CT757)
                                    mraY
                                             UDP-N-Acetylglucosamine Transferase
W
                  0571
                        (CT455)
                                    murA
                                              UDP-N-Acetylenolpyruvoylglucosamine Reductase
                                    murB
M
                  0988
                        (CT831)
                                    murC&ddlA Muramate-Ala Ligase & D-Ala-D-Alam Ligase
                        (CT762)
                  0905
                                              Muramoylalanine-Glutamate Ligase
                  0901
                        (CT758)
                                    murD
                                              N-Acetylmuramoylalanylglutamyl DAP Ligase
       30
                        (CT269)
                                    murE
                                              Muramoyl-DAP Ligase
                  0899
                        (CT756)
                                    murF
ļ.
                                              Peptidoglycan Transferase
                        (CT761)
                                    murG
                  0904
1=
                                              Muramidase (invasin repeat family)
                        (CT759)
                                    nlpD
                  0902
Ü
                                              PBP2-Transglycolase/Transpeptidase
                  0694
                        (CT682)
                                    pbp2
Ð
       35
                                              Transglycolase/Transpeptidase
                  0419
                        (CT270)
                                    pbp3
                        (CT272)
                                    yabC
                                              PBP2B Family Methyltransferase
                  0421
                                                                         Cellular Processes
                Cell Division
       40
                                              Axial Filament Protein
                  0959
                       (CT808)
                                    cafE
                                    ftsK
                                              Cell Division Protein FtsK
                  0880
                         (CT739)
                         (CT760)
                                    ftsW
                                              Cell Division Protein FtsW
                  0903
                                              Cell Division Protein FtsY
                  0972
                         (CT820)
                                    ftsY
                                              FAD-dependent Oxidoreductase
                   0617
                         (CT498)
                                    gidA
       45
                                              Chromosome Partitioning ATPase
                         (CT582)
                                    minD
                   0805
                   0850
                         (CT709)
                                    mreB
                                              Rod Shape Protein-Sugar Kinase
                   0867
                         (CT726)
                                    rodA
                                              Rod Shape Protein
                                              Chromosome Partitioning: Protein-
                  -0684 = (CT688) - parB=
                Detoxtification
       50
                                              Superoxide Dismutase (Mn)
                   0057 (CT294)
                                    sodM
                                              Thio-specific Antioxidant (TSA) Peroxidase
                   0778 (CT603)
                                    ahpC
                Signal Transduction
                                              S/T Protein Kinase
                   0148 (CT145)
                                              Two-Component Sensor
                   0584
                         (CT467)
                                    atoS
                                              cAMP-Dependent Protein Kinase Regulatory Subunit
       55
                   0294
                         (CT235)
                   0712 (CT664)
                                              (FHA domain)
```

KDO Transferase

0154 (CT208)

gseA

```
GTP Binding Protein
                                     hflX
                   0478 (CT379)
                                               S/T Protein Kinase
                   0703
                         (CT673)
                                               S/T Protein Kinase
                   0095
                         (CT301)
                   0397
                         (CT259)
                                               PP2C Phosphatase Family
         5
                                               PTS Phosphocarrier Protein Hpr
                   0037
                         (CT337)
                                     ptsH
                                               PTS PEP Phosphotransferase
                   0038
                         (CT336)
                                     ptsI
                                               PTS IIA Protein_1
                   0060
                         (CT291)
                                     ptsN_1
                                               PTS IIA Protein + HTH DNA-Binding Domain
                                     ptsN_2
                   0061
                         (CT290)
                                     surE
                                               SurE-like Acid Phosphatase
                   0262
                         (CT218)
                                               Thiophene/Furan Oxidation Protein
       10
                                     thdF
                   0838
                         (CT698)
                                               TPR Repeats-CT683 Hypothetical Protein
                   0693
                         (CT683)
                                               GTP Binding Protein
                         (CT092)
                                     ychF
                   0321
                                     yhbZ
                                               GTP binding protein
                   0544
                         (CT418)
                                               GTPase/GTP-binding protein
                         (CT703)
                                     yphC
                   0844
       15
                Standard Protein Secretion
                                               Signal Recognition Particle GTPase
                   0115
                         (CT025)
                                     ffh
                                               Flagellar Secretion Protein
                   0363
                         (CT060)
                                     flhA
                                               Flagellum-specific ATP Synthase
                   0858
                         (CT717)
                                     fliI
                                               Flagellar Motor Switch Domain/YscQ family
                   0704
                         (CT672)
                                     fliN
20
                                               Gen. Secretion Protein D
                         (CT572)
                                     gspD
                   0815
                                               Gen. Secretion Protein E
                   0816
                         (CT571)
                                     gspE
                                               Gen. Secretion Protein F
                   0817
                          (CT570)
                                      gspF
                                               GTPase
                   0359
                          (CT064)
                                      lepA
U
                                               Signal Peptidase I
                   0110
                          (CT020)
                                      lepB
       25
                                                Lipoprotein Signal Peptidase
ļ.
                   0535
                         (CT408)
                                     IspA
                                               Protein Translocase Subunit_l
W
                   0260
                         (CT141)
                                      secA 1
                                                Translocase SecA 2
U
                   0841
                          (CT701)
                                      secA_2
                                                  Protein Export Proteins SecD/SecF (fusion)
                   0564
                          (CT448)
                                      secD&secF
15
                                                Preprotein Translocase
                          (CT321)
                                      secE
                   0075
ļ.
       30
                                                Translocase
                   0629
                          (CT510)
                                      secY
ŀ÷
                                                Trigger Factor-Peptidyl-prolyl Isomerase
                   0848
                          (CT707)
                                      tig
ŀ÷
                 Transport-Related Proteins
ļ÷
                                                Hypothetical Proline Permease
                   0486
٠O
                                                Neutral Amino Acid (Glutamate) Transporter
                   0289
                          (CT230)
                                      aaaT
4D
       35
                                                ABC Transporter ATPase
                   0691
                          (CT685)
                                      abcX
                                                Arginine/Omithine Antiporter
                                      arcD
                    1031
                          (CT374)
                                                Arginine Periplasmic Binding Protein
                   0482
                          (CT381)
                                      arti
                                                Amino Acid (Branched) Transport
                   0836
                          (CT554)
                                      bmQ
                                                D-Ala/Gly Permease_I
                   0536
                           (CT409)
                                      dagA_1
        40
                                                D-Alanine/Glycine Permease_2
                    0876
                          (CT735)
                                      dagA_2
                                                ABC ATPase Dipeptide Transport
                    0682
                                      dppD
                          (CT690)
                                                ABC ATPase Dipeptide Transport
                    0683
                          (CT689)
                                      dppF
                                                Dipeptide Transporter ATPase
                    0280
                           (CT689)
                                      dppF
                           (CT596)
                                      exbB
                                                Macromolecule Transporter
                    0785
        45
                                                Biopolymer Transport Protein
                    0784
                           (CT597)
                                      exbD
                                      fliY
                                                Glutamine Binding Protein
                    0604
                           (CT486)
                                                ABC Amino Acid Transporter Permease
                    0192
                           (CT129)
                                      ginP
                                                ABC Amino Acid Transporter ATPase
                          _(CT130)_
                    0191_
                                      glnQ
                                                 Glutamate Symport
                    0528
                           (CT401)
                                      gltT
        50
                                      mgtE
                                                Mg++ Transporter (CBS Domain)
                    0286
                           (CT194)
                                                Transport ATP Binding Protein
                    0413
                           (CT264)
                                      msbA
                                                 Na+-dependent Transporter
                    0290
                           (CT231)
                                                 Oligopeptide Binding Protein_1
                    0195
                           (CT198)
                                      oppA 1
                                                Oligopeptide Binding Protein_2
                    0196
                           (CT198)
                                      oppA_2
        55
                                                 Oligopeptide Binding Protein_3
                           (CT139)
                    0198
                           (CT175)
                                      oppA_4
                                                Oligopeptide Binding Protein 4
```

```
0200
                         (CT200)
                                     oppC_1
                                               Oligopeptide Permease_1
         5
                                               Oligopeptide Permease_2
                   0597
                         (CT478)
                                     oppC_2
                                               Oligopeptide Transport ATPase
                   0201
                         (CT201)
                                     OppD
                                               Oligopeptide Transport ATPase
                   0202
                         (CT202)
                                     oppF
                   0231
                         (CT180)
                                     tauB
                                               ABC Transport ATPase (Nitrate/Fe)
                   0782
                         (CT599)
                                     tolB
                                               Macromolecule Transporter
       10
                                               Tyrosine Transport_1
                   0969
                         (CT817)
                                     tyrP_1
                                               Tyrosine Transport 2
                   0970
                                     tvrP 2
                         (CT818)
                   0665
                         (CT544)
                                     uhpC
                                               Hexosphosphate Transport
                   0282
                                               Amino Acid Transporter
                         (CT216)
                                     xasA
                   0207
                                     ybhl
                                               dicarboxylate Translocator
                         (CT204)
       15
                                               Transport Permease
                   0971
                         (CT819)
                                     yc¢A
                   0248
                         (CT152)
                                     ycfV
                                               ABC Transporter ATPase
                   1014
                         (CT856)
                                     ychM
                                               Sulfa e Transporter
                   0736
                         (CT641)
                                     ygeD
                                               Efflu: Protein
Phospitate Permease
                   0680
                         (CT692)
                                     ygo4
       20
                   0723
                         (CT653)
                                               ABC Transporter ATPase
                                     yhbG
                   0023
                         (CT348)
                                     ујјК
                                               ABC Transporter Protein ATPase
                                               Cationic Amino Acid Transporter
                   0127
                         (CT034)
                                     ytfF
                                               Solute Protein Binding Family
                   0349
                         (CT067)
                                     ytgA
                                               ABC transporter ATPase
ļ÷
                   0348
                         (CT068)
                                     ytgB
       25
Ü
                   0347
                         (CT069)
                                     ytgC
                                               Integral Membrane Protein
                                               Integral Membrane Protein
                   0346
                         (CT070)
                                     ytgD
LM
                                               ABC Transporter Permease
                   1012
                         (CT854)
                                     yzeB
                   0868
                         (CT727)
                                               Metal Transport P-type ATPase
                                     zntA
į
                   0279
                                               Possible ABC Transporter Permease Protein
14
       30
                   0543
                         (CT417)
                                               (Metal Transport Protein)
14
                   0692
                         (CT684)
                                               ABC Transporter
14
                                               ABC Transporter ATPase
                   0542
                         (CT416)
ŧŪ
                         (CT686)
                                               ABC Transporter Membrane Protein
                   0690
ıD
                   0541
                         (CT415)
                                               solute binding protein
       35
                Type-III Secretion
                   0323 .(CT090)
                                     lcrD
                                               Low Calcium Response D
                   0324
                         (CT089)
                                     IcrE
                                               Low Calcium Response E
                                               Low Ca Response Protein H_1
                   0811
                         (CT576)
                                     lcrH 1
                                               Low Calcium Response_2
                   1021
                         (CT862)
                                     lcrH_2
       40
                   0325
                         (CT088)
                                     sycE
                                               Secretion Chaperone
                   0702
                                               Yop C/Gen Secretion Protein D
                         (CT674)
                                     yscC
                   0828
                         (CT559)
                                               Yop Translocation J
                                     yscJ
                   0826
                         (CT561)
                                     yscL
                                               Yop Translocation L
                   0707
                         (CT669)
                                     yscN
                                               Yop N (Flagellar-Type ATPase)
       45
                   0825
                         (CT562)
                                    yscR
                                               Yop Translocation R
                   0824
                                    yscS
                                               YopS Translocation Protein
                         (CT563)
                   0823
                         (CT564)
                                     yscT
                                               YopT Tranlocation T
                   0322
                         (CT091)
                                     yscU
                                               Yop Translocation Protein U
       50
                                                                   Central Intermediary Metabolism
                Glycogen Metabolism
                   0856 (CT715)
                                               UDP-Glucose Pyrophosphorylase
                   0948
                         (CT798)
                                               Glycogen Synthase
                                     glgA
                                               Glucan Branching Enzyme
                   0475
                         (CT866)
                                     glgB
       55
                   0607
                         (CT489)
                                     glgC
                                               Glucose-1-P Adenyltransferase
                   0307
                         (CT248)
                                     glgP
                                               Glycogen Phosphorylase
```

Oligopeptide Binding Lipoprotein_5

Oligopeptide Permease_1

Oligopeptide Permease_2

0599

0199

0598

(CT480)

(CT199)

(CT479)

(CT042)

glgX

0388

oppA_5

oppB_I

oppB_2

Glycogen Hydrolase (debranching)

```
0326 (CT087)
                                             Glucanotransferase
                                   malO
                 0851 (CT710)
                                   pckA
                                             Phosphoenolpyruvate Carboxykinase
               Phosphorous & Sulfur
                 0548 (CT435)
                                   cysJ
                                             Sulfite Reductase
        5
                                             Sulfite Synthesis/Biphosphate Phosphatase
                 0920
                       (CT774)
                                   cysQ
                                             Sulphohydrolase
                 0025
                       (CT346)
                                   atsA
                                             Inorganic Pyrophosphatase
                 0918 (CT772)
                                   ppa
                                                       DNA Replication, Modification, Repair & Recombination
      10
               DNA Mismatch Repair
                                             3-Methyladenine DNA Glycosylase
                 0505
                                             DNA Mismatch Repair
                 0812
                        (CT575)
                                   mutL
                 0941
                        (CT792)
                                   mutS
                                             DNA Mismatch Repair
                                             Adenine Glycosylase
                 0402
                        (CT107)
                                   mutY
      15
                                             Endonuclease IV
                        (CT625)
                 0732
                                   nfo
                                             Enodnuclease III
                 0837
                        (CT697)
                                   nth
               DNA Modification
                 0596
                        (CT477)
                                   ada
                                             Methyltransferase
                        (CT024)
                                             A/G-specific Methylase
                 0114
                                   hemK
20
                        (CT748)
                                             Transcription-Repair Coupling
                 0891
                                   mfd
                                             Holliday Junction Helicase
                 0620
                        (CT501)
                                   ruvA
                                             Holliday Junction Helicase
                 0390
                        (CT040)
                                   ruvB
                        (CT502)
                                   ruvC
                                             Crossover Junction Endonuclease
                  0621
                                             Sms Protein
                  0053
                        (CT298)
                                   sms
ļ
      25
                                             Uracil DNA Glycosylase
                  0773
                        (CT607)
                                   ung
M
                                             Exodoxyribonuclease VII
                  1062
                        (CT329)
                                   xseA
N
               DNA Recombination
8
                  0762 (CT650)
                                   recA
                                             RecA Recombination Protein
-1
                  0738
                        (CT639)
                                   гесВ
                                             Exodeoxyribonuclease V, Beta
14
      30
                                             Exodeoxyribonuclease V, Gamma
                  0737
                        (CT640)
                                   recC
į.
                                   recD_1
                                             Exodeoxyribonuclease V (Alpha Subunit)_1
                 0123
                        (CT033)
₽4
                                             Exodeoxyribonuclease V, Alpha_2
                 0752
                        (CT652)
                                   recD_2
Ð
                  0339
                        (CT074)
                                   recF
                                             ABC Superfamily ATPase
Ð
                  0340
                        (CT074)
                                             (frame-shift with 0339)
      35
                        (CT447)
                 0563
                                             ssDNA Exonuclease
                                   recJ
                                             Recombination Protein
                 0299
                        (CT240)
                                   recR
               DNA Replication
                  0309
                        (CT250)
                                             Replication Initiation Protein_1
                                   dnaA_l
                  0424
                        (CT275)
                                   dnaA_2
                                             Replication Initiation Factor_2
      40
                  0616
                        (CT497)
                                   dnaB
                                             Replicative DNA Helicase
                                   dnaE
                                             DNA Pol III Alpha
                  0666
                        (CT545)
                        (CT794)
                                             DNA Primase
                  0942
                                   dnaG
                  0338
                        (CT075)
                                   dnaN
                                             DNA Pol III (Beta)
                  0410
                        (CT261)
                                   dnaQ_i
                                             DNA Pol III Epsilon Chain_1
      45
                                             DNA Pol III Epsilon Chain_2
                  0655
                        (CT536)
                                   dnaQ_2
                        (CT334)
                                             DNA Pol III Gamma and Tau_I
                  0040
                                   dnaX_i
                  0272
                        (CT187)
                                   dnaX_2
                                             DNA Pol III Gamma and Tau_2
                                             DNA Ligase
                 0149
                        (CT146)=
                                  dnlJ
                  0274
                        (CT189)
                                             DNA Gyrase Subunit A_1
                                   gytA_l
      50
                                   gyrA_2
                                             DNA Gyrase Subunit A_2
                  0716
                        (CT660)
                                             DNA Gyrase Subunit B_1
                  0275
                        (CT190)
                                   gyrB_1
                  0715
                        (CT661)
                                   gyrB_2
                                             DNA Gyrase Subunit B_2
                                             Integration Host Factor Alpha
                  0416
                        (CT267)
                                   himD
                  0612
                        (CT493)
                                   polA
                                             DNA Polymerase I
      55
                  0924
                        (CT778)
                                   priA
                                             Primosomal Protein N'
                  0386
                        (CT044)
                                              SS DNA Binding Protein
```

```
SWI/SNF family helicase 2
                 0849
                        (CT708)
                                             DNA Topoisomerase I-Fused to SWI Domain
                 0769
                        (CT643)
                                   topA
                 0024
                        (CT347)
                                   xerC
                                             Integrase/recombinase
       5
                                             Integrase/recombinase
                  1024
                       (CT864)
                                   xerD
               Eukaryotic-Type Chromatin Factors
                                             Histone-Like Developmental Protein
                 0886 (CT743)
                                   hctA
                                             Histone-like Protein 2
                 0384 (CT046)
                                   hctB
                        (CT737)
                                             SET Domain protein
      10
                                             SWIB (YM74) Complex Protein
                 0577
                       (CT460)
               UVR Exinuclease Repair System
                                             Excinuclease ABC Subunit A
                        (CT333)
                 0096
                                   uvrA
                                             Exinuclease ABC Subunit B
                                   uvrB
                 0801
                        (CT586)
                                             Excinuclease ABC, Subunit C
                  0940
                        (CT791)
                                   uvrC
      15
                  0772
                        (CT608)
                                   uvrD
                                             DNA Helicase
                                                                        Energy Metabolism
               Aerobic
                  0855
                        (CT714)
                                   gpdA
                                             Glycerol-3-P Dehydrogenase
      20
0743
                        (CT634)
                                             Ubiquinone Oxidoreductase, Alpha
                                   ngrA
                                             NADH (Ubiquinone) Dehydrogenase
                  0427
                        (CT278)
                                   ngr2
                                             NADH (Ubiquinone) Oxidoreductase, Gamma
                  0428
                        (CT279)
                                   nqr3
                  0429
                        (CT280)
                                   nqr4
                                             NADH (Ubiquinone) Reductase 4
Ø
                                             NADH (Ubiquinone) Reductase 5
                  0430
                        (CT281)
                                   nqr5
      25
                                             Phenolhydrolase/NADH (Ubiquinone) Oxidoreductase 6
ļ.,
                  0883
                        (CT740)
                                   nar6
Щ
               ATP Biogenesis and metabolism
                                              ADP/ATP Translocase_1
                  0351
                        (CT065)
                                   adt_l
M
                        (CT495)
                                   adt_2
                                              ADP/ATP Translocase_2
                  0614
                                              ATP Synthase Subunit A
                  0088
                        (CT308)
                                   atpA
14
      30
                                              ATP Synthase Subunit B
                        (CT307)
                  0089
                                   atpB
14
                        (CT306)
                                              ATP Synthase Subunit D
                  0090
                                   atpD
}#
                  0086
                        (CT310)
                                   aφE
                                              ATP Synthase Subunit E
ATP Synthase Subunit I
                  0091
                        (CT305)
                                   atpl
J
                        (CT304)
                                   atpK
                                              ATP Synthase Subunit K
                  0092
Ð
      35
                        (CT719)
                                   fliF
                                              Flagellar M-Ring Protein
                  0860
               Electron Transport Chain
                  0102 (CT013)
                                   cydA
                                              Cytochrome Oxidase Subunit I
                                             Cytochrome Oxidase Subunit II
                  0103 (CT014)
                                   cydB
                  0364 (CT059)
                                              Ferredoxin
      40
                                              Predicted Ferredoxin
                  0084 (CT312)
               Glycolysis & Gluconeogenesis
                                              Predicted 1,6-Fructose Biphosphate Aldolase
                  0281
                        (CT215)
                                   dhnA
                                              Enolase
                  0800
                        (CT587)
                                    eno
                  0624
                        (CT505)
                                    gapA
                                              Glyceraldehyde-3-P Dehyrogenase
      45
                  0056
                        (CT295)
                                              Phosphomannomutase
                                    mrsA
                  0967
                        (CT815)
                                              Phosphoglucomutase
                                    pgm
                                              Fructose-6-P Phosphotransferase_1
                  0160
                        (CT207)
                                    pfkA_I
                  0208
                        =(CT205)=
                                    pfkA_2
                                              Fructose=6-P Phosphotransferase_2
                        (CT378)
                                              Glucose-6-P Isomerase
                  1025
                                    pgi
       50
                                              Phosphoglycerate Kinase
                  0679
                        (CT693)
                                    pgk
                  0863
                        (CT722)
                                    pgmA
                                              Phosphoglycerate Mutase
                  0097
                        (CT332)
                                    pyk
                                              Pyruvate Kinase
                                              Triosephosphate Isomerase
                  1063
                        (CT328)
                                    tpiS
               Pentose Phosphate Pathway
       55
                                              Glucose-6-P Dehyrogenase (DevB family)
                  0239 (CT186)
                                    devB
                                              Transketolase
                  1060 (CT331)
                                    dxs
```

SWI/SNF family helicase_1

0835

(CT555)

```
Ribulose-P Epimerase
                  0185
                         (CT121)
                                    тре
                         (CT213)
                                              Ribose-5-P Isomerase A
                                    тріA
                  0141
                                              Transaldolase
                  0083
                         (CT313)
                                    tal
        5
                  0893
                         (CT750)
                                    tktB
                                              Transketolase
                                              Glucose-6-P Dehyrogenase
                  0238
                        (CT185)
                                    zwf
               Pyruvate Dehydrogenase
                        (CT557)
                                    Abal
                                              Lipoamide Dehydrogenase
                  0833
                                              Lipoate Protein Ligase-Like Protein
                        (CT285)
                                    \mathsf{lplA}_{-}\mathsf{l}
                  0436
      10
                  0618
                         (CT499)
                                    lplA_2
                                              Lipoate-Protein Ligase A
                                    pdhA&B
                                              Oxoisovalerate Dehydrogenase α/β Fusion
                  0033
                        (CT340)
                                    pdhA
                                              Pyruvate Dehydrogenase Alpha
                  0304
                         (CT245)
                                              Pyruvate Dehydrogenase Beta
                  0305
                         (CT246)
                                    pdhB
                        (CT247)
                                    pdhC
                                              Dihydrolipoamide Acetyltransferase
                  0306
      15
               TCA Cycle
                  0495
                       (CT390)
                                    aspC
                                              Aspartate Aminotransferase
                                    fumC
                  1013
                         (CT855)
                                              Furnarate Hydratase
                  1028
                         (CT376)
                                    mdhC
                                              Malate Dehyrogenase
                  0789
                         (CT592)
                                    sdhA
                                              Succinate Dehydrogenase
20
                  0790
                         (CT591)
                                    sdhB
                                              Succinate Dehydrogenase
                                              Succinate Dehydrogenase
                                    sdhC
                  0788
                        (CT593)
                                              Oxoglutarate Dehydrogenase
                  0378
                         (CT054)
                                    sucA
                  0377
                         (CT055)
                                    sucB_1
                                              Dihydrolipoamide Succinyltransferase_I
                                              Dihydrolipoamide Succinyltransferase_2
                  0527
                         (CT400)
                                    sucB_2
      25
                         (CT821)
                                    sucC
                                              Succinyl-CoA Synthetase, Beta
                  0973
                                              Succinyl-CoA Synthetase, Alpha
                  0974
                         (CT822)
                                    sucD
M
#
                                                               Protein Folding, Assembly & Modification
14
                Chaperones
ļ.
      30
                  0949 (CT799)
                                              General Stress Protein
                                    ctc
ļ.
                  0534
                        (CT407)
                                    dksA
                                              DnaK Suppressor
ļ.
                  0032
                         (CT341)
                                    dnaJ
                                              Heat Shock Protein J
Q
                         (CT396)
                                              Hsp-70
                  0503
                                    dnaK
Ď
                         (CT110)
                                              Hsp-60_1
                  0134
                                    groEL 1
      35
                  0777
                         (CT604)
                                    groEL_2
                                              Hsp-60_2
                  0898
                         (CT755)
                                    groEL_3
                                              Hsp-60_3
                  0135
                         (CT111)
                                    groES
                                              10KDa Chaperonin
                  0502
                         (CT395)
                                    grpE
                                              HSP-70 Cofactor
                         (CT541)
                                              FKBP-type Peptidyl-prolyl Cis-Trans Isomerase
                  0661
                                    mip
      40
               Proteases
                  0144
                         (CT113)
                                    clpB
                                              Clp Protease ATPase
                  0437
                         (CT286)
                                    clpC
                                              ClpC Protease
                         (CT431)
                                    clpP_1
                                              CLP Protease
                  0520
                         (CT706)
                                    clpP_2
                                              CLP Protease Subunit
                  0847
      45
                                              CLP Protease ATPase
                         (CT705)
                  0846
                                    clpX
                  0269
                         (CT138)
                                              Dipeptidase
                  0998
                         (CT841)
                                    ftsH
                                              ATP-dependent Zinc Protease
                  0030
                         (CT343)
                                              O-Sialoglycoprotein Endopeptidase_l-
                                    gcp_1-
                                              O-Sialoglycoprotein Endopeptidase_2
                         (CT197)
                  0194
                                    gcp_2
      50
                  0979
                         (CT823)
                                    htrA
                                              DO Serine Protease
                                              Insulinase family/Protease III
                  0957
                         (CT806)
                                    ide
                                              Lon ATP-dependent Protease
                  0027
                         (CT344)
                                    lon
                  1017
                         (CT859)
                                    lytB
                                              Metalloprotease
                         (CT851)
                                              Methionine Aminopeptidase
                                    map
       55
                                              Leucyl Aminopeptidase A
                  0385
                         (CT045)
                                    pepA
```

Oligopeptidase

6-Phosphogluconate Dehydrogenase

(CT063)

0360

0136

(CT112)

pepF

	0813	(CT574)	pepP	Aminopeptidase P
	0613	(CT494)	sohB	Protease
	0555	(CT441)	tsp	Tail-Specific Protease
	0344	(CT072)	yaeL	Metalloprotease
5	1890	(CT824)		Zinc Metalloprotease (insulinase family)
	Protein Is	omerases		
	0227	(CT176)	dsbB	Disulfide bond Oxidoreductase
	0786	(CT595)	dsbD	Thio:disulfide Interchange Protein
	0228	(CT177)	dsbG	Disulfide Bond Chaperone
10	0933	(CT783)		Predicted Disulfide Bond Isomerase
	0926	(CT780)		Thioredoxin Disulfide Isomerase

Transcription

```
RNA Degradation
                                             Polyribonucleotide Nucleotidyltransferase
                  0999 (CT842)
        5
                                             Ribonuclease III
                  0054
                        (CT297)
                                             Ribonuclease HII_1
                 0119
                        (CT029)
                                   mhB 1
                                             Ribonuclease HII 2
                        (CT008)
                  1068
                                   mhB 2
                                              Ribonuclease P Protein Component
                  0934
                        (CT784)
                                    mpA
                  0504
                        (CT397)
                                    vacB
                                              Ribonuclease Family
      10
               RNA Elongation & Termination Factors
                  0741
                        (CT636)
                                    greA
                                             Transcription Elongation Factor
                        (CT097)
                                             N Utilization Protein A
                  0316
                                    nusA
                                             Transcriptional Antitermination
                  0076
                        (CT320)
                                    nusG
                  0845
                         (CT704)
                                    pcnB_1
                                             Poly A Polymerase_!
      15
                  0966
                        (CT410)
                                    pcnB_2
                                             PolyA Polymerase_2
                  0610 (CT491)
                                              Transcription Termination Factor
                                    rho
               RNA Methylases
                                              RNA Methyltransferase
                  0674 (CT553)
                                    fmu
                  1059
                        (CT354)
                                    kgsA
                                              Dimethyladenosine Transferase
20
                  0187
                         (CT133)
                                              Predicted Methylase
                  0530
                        (CT403)
                                    spoU_I
                                              rRNA Methylase_I
                        (CT540)
                                    spoU 2
                                              rRNA Methylase_2
                  0660
                                              tRNA (Guanine N-1)-Methyltransferase
                                    tmD
                  0117
                        (CT027)
Ü
                  0885
                        (CT742)
                                    ygcA
                                              rRNA Methyltransferse
      25
                        (CT829)
                                              Predicted rRNA Methylase
ļ.
                  0986
                                    yggH
Œ
                  0987
                        (CT830)
                                    ytgB
                                              Predicted rRNA Methylase
M
               RNA Modification
                                              Methionyl tRNA Formyltransferase
                  0649 (CT530)
                                    fmt
                                              tRNA Pyrophosphate Transferase
                  0910
                        (CT766)
                                    miaA
      30
                                              Predicted Pseudouridine Synthase
                  0719
                         (CT658)
                                    sfhB
ļ.
                  0219
                         (CT193)
                                              Queuine tRNA Ribosyl Transferase
                                    tgt
1
                                              Pseudouridylate Synthase I
                  0580
                        (CT463)
                                    truA
1
                                              tRNA Pseudouridine Synthase
                  0319
                       (CT094)
                                    truB
G
                                    yceC
                                              Predicted Pseudouridine Synthetase Family
                         (CT106)
                  0403
٠D
      35
                                              Predicted Pseudouridine Synthase
                  0864
                         (CT723)
                                    yjbC
                RNA Polymerase & Transcription Regulators
                  0586 (CT468)
                                    atoC
                                              Two-Component Regulator
                  0362
                         (CT061)
                                    rpsD
                                              Sigma-28/WhiG Family
                  0501
                         (CT394)
                                              HTH Transcriptional Repressor
                                    hrcA
      40
                                              Sigma Regulatory Family Protein-PP2C Phosphatase (RsbW Antagonist)
                  0793
                         (CT588)
                                    rbsU
                                              RNA Polymerase Alpha
                  0626
                         (CT507)
                                    фοΑ
                  0081
                         (CT315)
                                    троВ
                                              RNA Polymerase Beta
                         (CT314)
                                              RNA Polymerase Beta'
                  0082
                                    гроС
                  0756
                         (CT615)
                                              RNA Polymerase Sigma-66
                                    тоО
       45
                                              RNA Polymerase Sigma-54
                  0771
                         (CT609)
                                    τροΝ
                  0511
                         (CT424)
                                    rsbV_1
                                              Sigma Regulatory Factor_1
                         (CT765)
                                              Sigma Factor Regulator_2
                                    rsbV_2
                                              Sigma-Regulatory-Factor-Histidine-Kinase ---
                  0670 (CT549)
                                    rsbW:
                                              HTH Transcriptional Regulatory Protein + Receiver Doman
                  0750
                        (CT630)
                                    tctD
       50
                                              HTH Transcriptional Regulator
                   1069
                        (CT009)
                                    yfgA
                                                                            Translation
                Amino Acyl tRNA Synthesis
                  0892 (CT749)
                                    alaS
                                              Alanyl tRNA Synthetase
       55
```

Arginyl tRNA Transferase

Aspartyl tRNA Synthetase

0570

(CT454)

0662 (CT542)

argS aspS

```
0932
                           (CT782)
                                      cysS
                                                Cysteinyl tRNA Synthetase
                           (CT003)
                    0003
                                                Glu tRNA Gln Amidotransferase (A subunit)
                                      gatA
                    0004
                           (CT004)
                                      gatB
                                                Glu tRNA Gln Amidotransferase (B Subunit)
                                      gatC
                                                Glu tRNA Gln Amidotransferase (C subunit)
                    0002
                           (CT002)
          5
                    0560
                           (CT445)
                                      gltX
                                                Glutamyl-tRNA Synthetase
                    0946
                          (CT796)
                                                Glycyl tRNA Synthetase
                                      glyQ
                    0663
                          (CT543)
                                      hisS
                                                Histidyl tRNA Synthetase
                    0109
                                                Isoleucyl-tRNA Synthetase
                          (CT019)
                                      ileS
                    0153
                          (CT209)
                                      leuS
                                                Leucyl tRNA Synthetase
        10
                    0931
                          (CT781)
                                      lysS
                                                Lysyl tRNA Synthetase
                    0122
                          (CT032)
                                      metG
                                                Methionyl-tRNA Synthetase
                                      pheS
                    0993
                          (CT836)
                                                Phenylalanyl tRNA Synthetase, Alpha
                    0594
                          (CT475)
                                      pheT
                                                Phenylalanyl tRNA Synthetase Beta
                    0500
                          (CT393)
                                      proS
                                                Prolyi tRNA Synthetase
        15
                    0870
                          (CT729)
                                      ser$
                                                Seryl tRNA Synthetase 2
                          (CT581)
                    0806
                                      thrS
                                                Threonyl tRNA Synthetase
                    0802
                          (CT585)
                                      trpS
                                                Tryptophanyl tRNA Synthetase
                    0361
                          (CT062)
                                      tytS
                                                Tyrosyl tRNA Synthetase
                    0094
                          (CT302)
                                      valS
                                                Valyl tRNA Synthetase
       20
                 Peptide Chain Initiation, Elongation & Termination
1067
                          (CT353)
                                     def
                                                Polypeptide Deformylase
                   0184
                          (CT122)
                                      cfp_1
                                                Elongation Factor P_1
                   0895
                          (CT752)
                                      efp_2
                                                Elongation Factor P_2
                   0550
                          (CT437)
                                                Elongation Factor G
                                      fusA
       25
                   0073
                          (CT323)
                                                Initiation Factor IF-1
                                     infA
Ø
                                     infB
                   0317
                          (CT096)
                                                Initiation Factor-2
101
                   0990
                          (CT833)
                                     infC
                                                Initiation Factor 3
                          (CT023)
                                     pfrA
                                                Peptide Chain Releasing Factor 1
Ħ
                   0576
                          (CT459)
                                                Peptide Chain Release Factor 2
-
                                     prfB
       30
                   0950
                          (CT800)
                                                Peptidyl tRNA Hydrolase
14
                                     pth
                   0318
                          (CT095)
                                     rbfА
                                                Ribosome Binding Factor A
ļ±
                   0699
                          (CT677)
                                     пſ
                                                Ribosome Releasing Factor
ŀ÷
                   0697
                          (CT679)
                                     tsf
                                                Elongation Factor TS
Đ
                   0074
                         (CT322)
                                     tufA
                                                Elongation Factor Tu
Ü
       35
                 Ribosomal Proteins
                   0078 · (CT318)
                                     rl1
                                                L1 Ribosomal Protein
                   0644
                          (CT525)
                                     rI2
                                                L2 Ribosomal Protein
                   0647
                          (CT528)
                                     rl3
                                                L3 Ribosomal Protein
                   0646
                          (CT527)
                                     rl4
                                                L4 Ribosomal Protein
       40
                   0635
                          (CT516)
                                     rl5
                                                L5 Ribosomal Protein
                   0633
                          (CT514)
                                     rl6
                                                L6 Ribosomal Protein
                   0080
                          (CT316)
                                     rl7
                                                L7/L12 Ribosomal Protein
                   0953
                          (CT803)
                                     rl9
                                                L9 Ribosomal Protein
                   0079
                          (CT317)
                                     rllO
                                               L10 Ribosomal Protein
       45
                   0077
                          (CT319)
                                     rll I
                                               L11 Ribosomal Protein
                   0247
                          (CT125)
                                     rl13
                                               L13 Ribosomal Protein
                   0637
                          (CT518)
                                     rl14
                                               L14 Ribosomal Protein
                   0630
                          (CT511).
                                     _rl15__
                                               L15 Ribosomal Protein
                          (CT521)
                   0640
                                     rl16
                                               L16 Ribosomal Protein
       50
                   0625
                          (CT506)
                                     rl17
                                               L17 Ribosomal Protein
                          (CT513)
                   0632
                                     rl18
                                               L18 Ribosomal Protein
                   0118
                          (CT028)
                                     rl19
                                               L19 Ribosomal Protein
                   0992
                          (CT835)
                                     r120
                                               L20 Ribosomal Protein
                   0546
                          (CT420)
                                     rl21
                                               L21 Ribosomal Protein
       55
                         (CT523)
                                     rl22
                   0642
                                               L22 Ribosomal Protein
                         (CT526)
                                     rl23
                   0645
                                               L23 Ribosomal Protein
```

```
0545
                         (CT419)
                                    r127
                                              L27 ribosomal protein
                  0327
                         (CT086)
                                    rl28
                                              L28 Ribosomal Protein
                         (CT520)
                                    rl29
                                              L29 Ribosomal Protein
                  0639
       5
                         (CT022)
                                    rl31
                                              L31 Ribosomal Protein
                  0112
                  0961
                         (CT810)
                                    rl32
                                              L32 Ribosomal Protein
                  0250
                         (CT150)
                                    rl33
                                              L33 Ribosomal Protein
                         (CT785)
                                    rl34
                                              L34 Ribosomal Protein
                  0935
                                    rl35
                                              L35 Ribosomal Protein
                  0991
                         (CT834)
      10
                        (CT786)
                                    rl36
                                              L36 Ribosomal Protein
                  0936
                  0315
                         (CT098)
                                    rs l
                                              S1 Ribosomal Protein
                        (CT680)
                                    rs2
                                              S2 Ribosomal Protein
                  0696
                                              S3 Ribosomal Protein
                        (CT522)
                                    rs3
                 0641
                  0733
                        (CT626)
                                    rs4
                                              S4 Ribosomal Protein
     15
                  0631
                        (CT512)
                                    rs5
                                              S5 Ribosomal Protein
                        (CT801)
                                              S6 Ribosomal Protein
                  0951
                                    rs6
                 0551
                        (CT438)
                                    rs7
                                              S7 Ribosomal Protein
                  0634
                        (CT515)
                                              S8 Ribosomal Protein
                                    rs8
                  0246
                        (CT126)
                                    rs9
                                              S9 Ribosomal Protein
     20
                  0549
                        (CT436)
                                    rs I O
                                              S10 Ribosomal Protein
0627
                        (CT508)
                                    rs I I
                                              S11 Ribosomal Protein
                 0552
                        (CT439)
                                    rs12
                                              S12 Ribosomal Protein
                  0628
                        (CT509)
                                    rs13
                                              S13 Ribosomal Protein
                  0937
                        (CT787)
                                    rs14
                                              S14 Ribosomal Protein
     25
                  1000
                        (CT843)
                                    rs 15
                                              S15 Ribosomal Protein
                 0116
                        (CT026)
                                    rs16
                                              S16 Ribosomal Protein
Ø
                 0638
                        (CT519)
                                    rs 17
                                              S17 Ribosomal Protein
IT
                 0952
                        (CT802)
                                    rs18
                                              S18 Ribosomal Protein
                        (CT524)
                 0643
                                    rs 19
                                              S19 Ribosomal Protein
14
     30
                 0754
                        (CT617)
                                    rs20
                                              S20 Ribosomal Protein
į mis
                 0031
                        (CT342)
                                    rs21
                                              S21 Ribosomal Protein
ļ
|=
Ū
     35
O
                                                                          Other Categories
              Chlamydia-Specific Proteins
                 0561
                        (CT446)
                                             CHLPS Euo Protein
                 0804
                        (CT583)
                                   Gp6D
                                             CHLTR Plasmid Paralog-
                        (CT119)
                                             Similarity to IncA_1
                 0186
     40
                 0291
                        (CT232)
                                   incB
                                             Inclusion Membrane Protein B
                 0292
                        (CT233)
                                   incC
                                             Inclusion Membrane Protein C
                                             LtuA Protein
                 1026
                        (CT377)
                 0333
                        (CT080)
                                             LtuB Protein
                 0005
                        (CT871)
                                   pmp I
                                             Polymorphic Outer Membrane Protein G Family
     45
                 0013
                        (CT871)
                                   pmp_2
                                             Polymorphic Outer Membrane Protein G Family
                 0014
                        (CT871)
                                             Polymorphic Outer Membrane Protein G Family
                                   pmp_3
                 0015
                        (CT871)
                                             PMP 3 (frame-shift with 0014)
                                   pmp_3
                 0016
                        (CT874)
                                   pmp_4
                                             Polymorphic Outer Membrane Protein G Family
                 0017
                        (CT871)
                                   pmp_4
                                             PMP_4 (frame-shift with 0016)
     50
                 0018
                        (CT874)
                                   pmp_5
                                             Polymorphic Outer Membrane Protein G Family
                 0019
                        (CT871)
                                             PMP_5 (frame-shift with 0018)
                                   pmp_5
                 0444
                        (CT871)
                                   pmp_6
                                             Polymorphic Outer Membrane Protein G/I Family
                 0445
                        (CT871)
                                   pmp_7
                                             Polymorphic Outer Membrane Protein G Family
                 0446
                        (CT871)
                                   pmp_8
                                             Polymorphic Outer Membrane Protein G Family
     55
                        (CT871)
                                             Polymorphic Outer Membrane Protein G/I Family
                 0447
                                   pmp 9
                        (CT871)
                                             Polymorphic Outer Membrane Protein G Family
                 0450
                                   pmp_10
                       (CT871)
                                             PMP_10 (Frame-shift with 0450)
```

pmp_10

0636 (CT517)

rl24

L24 Ribosomal Protein

```
Polymorphic Outer Membrane Protein (truncated) A/I Family
                  0452
                         (CT874)
                                    pmp_12
                  0453
                         (CT871)
                                    pmp_13
                                              Polymorphic Outer Membrane Protein G Family
                                              Polymorphic Outer Membrane Protein H Family
                         (CT872)
                                    pmp_14
        5
                                    pmp_15
                                              Polymorphic Outer Membrane Protein E Family
                         (CT869)
                  0466
                                              Polymorphic Outer Membrane Protein E Family
                  0467
                         (CT869)
                                    pmp_16
                  0468
                         (CT869)
                                    pmp_17
                                              Polymorphic Outer Membrane Protein E Family
                                              PMP_17 (Frame-shift with 0468)
                  0469
                         (CT869)
                                    pmp_17
                                    pmp_17
                                              PMP_17 (Frame-shift with 0469)
                  0470
                         (CT869)
      10
                         (CT870)
                                              Polymorphic Outer Membrane Protein E/F Family
                  0471
                                    pmp_18
                  0539
                         (CT412)
                                    pmp_19
                                              Polymorphic Membrane Protein A Family
                         (CT413)
                  0540
                                    pmp_20
                                              Polymorphic Membrane Protein B Family
                        (CT812)
                  0963
                                              Polymorphic Membrane Protein D Family
                                    pmp_21
                  0562
                                              CHLPS 43 kDa Protein Homolog 1
      15
                  0927
                                              CHLPS 43 kDa Protein Homolog_2
                  0928
                                              CHL?S 43 kDa Protein Homolog_3
                  0929
                                              CHL2S 43 kDa Protein Homolog 4
                  0728
                        (CT622)
                                              CHLPN 76kDa Homolog_I (CT622)
                  0729
                         (CT623)
                                              CHLPN 76kDa Homolog_2 (CT623)
      20
0133
                         (CT109)
                                              CHLPS Hypothetical Protein
0332
                        (CT081)
                                              CHL"R T2 Protein
               Miscellaneous Enzymes/Conserved Proxins
                  0193
                                   argR
                                              Possi le Arginine Repressor
Ü
                  1046
                                              Aron atic Amino Acid Hydroxylase
      25
-
                  0232
                                              Similarity to 5'-Methylthioadenosine Nucleosidase
                  0128
                        (CT035)
H
                                              Biotin Protein Ligase
                  0513
                        (CT426)
                                              Fe-S Oxidoreductase_1
M
                  0911
                        (CT767)
                                              Fe-S Oxidoreductase 2
                  0373
                        (CT057)
                                   gcpE
                                             GcpE Protein
ļ.b
      30
                  0407
                        (CT103)
                                             HAD Superfamily Hydrolase/Phosphatase
ŀ
                  0917
                        (CT771)
                                             Hydrolase/Phosphatase Homolog
ļш
                  0488
                        (CT385)
                                   ycfF
                                             HIT Family Hydrolase
ļњ
                  0701
                        (CT675)
                                   karG
                                             Arginine Kinase
Ð
                  0526
                        (CT399)
                                   kpsF
                                              GutQ/KpsF Family Sugar-P Isomerase
Ū
      35
                  0919
                        (CT773)
                                   ldh
                                             Leucine Dehydrogenase
                  0022
                        (CT349)
                                   maf
                                             Maf protein
                 0997
                        (CT840)
                                   mesJ
                                             PP-loop superfamily ATPase
                  0151
                        (CT148)
                                   mhpA
                                             Monooxygenase
                 0730
                        (CT624)
                                             Integral Membrane Protein
      40
                 0861
                        (CT720)
                                             NifU-Related Protein
                 0479
                        (CT380)
                                   ohnP
                                             Metal Dependent Hydrolase
                 0106
                        (CT015)
                                   phoH
                                             ATPase
                 0329
                        (CT084)
                                             Phopholipase D Superfamily
                        (CT284)
                 0435
                                             Phospholipase D Superfamily
      45
                 0581
                        (CT464)
                                             Phosphoglycolate Phosphatase
                 0897
                        (CT754)
                                             Predicted Phosphohydrolase
                 0509
                        (CT422)
                                             Predicted Metalloenzyme
                  1030
                        (CT375)
                                             Predicted D-Amino Acid Dehyrogenase
                 0531
                        (CT404)
                                             SAM Dependent Methyltransferase
     50
                 0337
                        (CT076)
                                   smpB
                                             Small Protein B
                 0394
                        (CT256)
                                   tlyC_l
                                             CBS Domain Protein (Hemolysin Homolog)_1
                 0510
                        (CT423)
                                   tlyC_2
                                             CBS Domains (Hemolysin Homolog)_2
                 0382
                        (CT048)
                                   yabC
                                             SAM-Dependent Methytransferase
                 0787
                        (CT594)
                                   yabD
                                             PHP Superfamily (Urease/Pyrimidinase) Hydrolase
     55
                 0611
                        (CT492)
                                   yacE
                                             Predicted Phosphatase/Kinase
                 0579
                        (CT462)
                                   yacM
                                             Sugar Nucleotide Phosphorylase
                 0578
                       (CT461)
                                   yael
                                             Phosphohydrolase
```

Polymorphic Outer Membrane Protein G Family

0451

(CT871)

pmp 11

```
CT071 Hypothetical Protein
                         (CT071)
                  0345
                                    yaeM
                  0566
                         (CT450)
                                    yaeS
                                               YaeS family Hypothetical Protein
                                               YagE family
                  0591
                         (CT472)
                                    yagE
                                    ybaB
                                               YbaB family Hypothetical Protein
                  0039
                         (CT335)
        5
                                    ybbP
                                               YbbP family Hypothetical Protein
                         (CT012)
                  0101
                  0915
                         (CT769)
                                    ybeB
                                               iojap Superfamily Ortholog
                         (CT108)
                                               ACR family
                  0137
                                    ybgi
                  0529
                         (CT402)
                                    ycaH
                                              ATPase
                                              PP-loop Superfamily ATPase
                         (CT287)
                                    ycbF
                  0438
      10
                         (CT627)
                                              YceA Hypothetical Protein
                  0734
                                    yceA
                                    ychB
                  0954
                         (CT804)
                                               Predicted Kinase
                  0261
                         (CT217)
                                    ydaO
                                              PP-Loop Superfamily ATPase
                                    ydhO
                  0245
                         (CT127)
                                              Polysaccharide Hydrolase-Invasin Repeat Family
                  0573
                         (CT457)
                                    yebC
                                               YebC Family Hypothetical Protein
      15
                  0689
                         (CT687)
                                    yfhO_1
                                              NifS-related Aminotransferase_1
                  0862
                         (CT721)
                                    yfhO_2
                                              NifS-related Aminotransferase_2
                  0547
                         (CT434)
                                               YgbB Family Hypothetical Protein
                                    ygbB
                  0237
                         (CT184)
                                    yggF
                                               YggF Family Hypothetical Protein
                  0775
                         (CT606)
                                    ygg∨
                                               YggV Family Hypothetical Protein
      20
                  0396
                         (CT258)
                                    yhfO_3
                                              NifS-related Aminotransferase_3
0605
                         (CT487)
                                    yhhF
                                              Predicted Methylase
                  0575
                         (CT458)
                                    yhhY
                                              Amino Group Acetyl Transferase
                  0592
                         (CT473)
                                    yidD
                                              YidD Family
                  0982
                         (CT825)
                                    yigN
                                              YigN Family Hypothetical Protein
      25
                  0657
                         (CT537)
                                    yjeE
                                              YjeE Hypothetical Protein
                 0768
                         (CT644)
                                              Yohl Predicted Oxidoreductase
                                    yohl
ij
                  0336
                        (CT077)
                                    yojL
                                              YojL Hypothetical Protein
H
                  0217
                         (CT140)
                                    ypdP
                                              YpdP Hypothetical Protein
ij
                         (CT212)
                  0140
                                    yqdE
                                              YqdE Hypothetical Protein
1-4
      30
                        (CT221)
                  0263
                                    yqfU
                                              YqfU Hypothetical Protein
14
                 0139
                        (CT211)
                                    yqgE
                                              YqgE Hypothetical Protein
[=
                  0270
                         (CT137)
                                    ywlC
                                              SuA5 Superfamily-related Protein
ŀ÷
                         (CT738)
                  0879
                                    yycJ
                                              Metal Dependent Hydrolase
ű
     35
                                                            Homologs to CHLTR Hypothetical Coding Genes
                        (CT001)
                  0001
                                    CT001 Hypothetical Protein
                  0020
                         (CT351)
                                    CT351 Hypothetical Protein
                        (CT350)
                  0021
                                    CT350 Hypothetical Protein
                        (CT345)
                  0026
                                    CT345 Hypothetical Protein
     40
                 0035
                        (CT339)
                                    CT339 Hypothetical Protein
                 0036
                        (CT338)
                                    CT338 Hypothetical Protein
                 0055
                        (CT296)
                                    CT296 Hypothetical Protein
                 0062
                        (CT289)
                                    CT289 Hypothetical Protein
                        (CT288)
                 0065
                                    CT288 Hypothetical Protein
     45
                 0068
                        (CT360)
                                    CT360 Hypothetical Protein
                 0071
                        (CT325)
                                    CT325 Hypothetical Protein
                 0072
                        (CT324)
                                    CT324 Hypothetical Protein
                        (CT311)
                 0085
                                    CT311 Hypothetical Protein
                 0087
                        (CT309)
                                    CT309 Hypothetical Protein
     50
                 0093
                        (CT303)
                                    CT303 Hypothetical Protein
                 0100
                        (CT011)
                                    CT011 Hypothetical Protein
                 0104
                        (CT017)
                                    CT017 Hypothetical Protein
                 0105
                        (CT016)
                                    CT016 Hypothetical Protein
                        (CT058)
                                    CT058 Hypothetical Protein_1
                 0107
     55
                        (CT018)
                 0108
                                    CT018 Similarity
                 0111
                        (CT021)
                                    CT021 Hypothetical Protein
                        (CT031)
                                    CT031 Hypothetical Protein
                 0121
```

```
0129
                         (CT036)
                                     CT036 Similarity
                         (CT114)
                                     CT114 Hypothetical Protein
                   0145
                         (CT147)
                                     CT147 Hypothetical Protein
                   0150
                         (CT149)
                                     CT149 Hypothetical Protein
                   0152
         5
                         (CT153)
                                     CT153 Hypothetical Protein
                   0176
                                     CT132 Hypothetical Protein
                   0188
                         (CT132)
                   0189
                         (CT131)
                                     CT131 Hypothetical Protein
                                     CT203 Hypothetical Protein
                   0206
                         (CT203)
                                     CT178 Hypothetical Protein
                   0229
                         (CT178)
                                     CT179 Hypothetical Protein
       10
                   0230
                         (CT179)
                                     CT181 Hypothetical Protein
                   0234
                          (CT181)
                                     CT151 Hypothetical Protein
                   0249
                          (CT151)
                                     CT144 Hypothetical Protein_1
                   0253
                          (CT144)
                                     CT143 Hypothetical Protein_I
                   0254
                          (CT143)
       15
                                     CT142 Hypothetical Protein_1
                   0255
                          (CT142)
                          (CT144)
                                     CT144 Hypothetical Protein_2
                   0256
                                     CT143 Hypothetical Protein_2
                          (CT143)
                   0257
                                     CT142 Hypothetical Protein_2
                          (CT142)
                   0259
                                     CT191 Hypothetical Protein
                   0276
                         (CT191)
       20
                                     CT195 Hypothetical Protein
                   0288
                          (CT195)
CT234 Hypothetical Protein
                   0293
                          (CT234)
                                     CT368 Hypothetical Protein
                          (CT242)
                   0301
                                     CT244 Hypothetical Protein
                          (CT244)
                   0303
                                     CT249 Similarity
                   0308
                          (CT249)
       25
                   0312
                          (CT101)
                                     CT101 Hypothetical Protein
                                     CT085 Hypothetical Protein
                   0328
                          (CT085)
                                     CT083 Hypothetical Protein
                          (CT083)
                   0330
                                     CT082 Hypothetical Protein
                          (CT082)
                   0331
                                     CT079 Similarity
                   0334
                          (CT079)
ŀ÷
       30
                          (CT073)
                                     CT073 Hypothetical Protein
                   0342
ŀ÷
                                     (frame-shift with 0342?)
                   0343
                          (CT073)
Į÷.
                                     CT066 Hypothetical Protein
                   0350
                          (CT066)
i de
                          (CT058)
                                     CT058 Hypothetical Protein_2
                   0369
J
                                     CT058 Hypothetical Protein_3
                   0370
                          (CT058)
4D
       35
                          (CT056)
                                     CT056 Hypothetical Protein
                   0374
                   0379
                          (CT053)
                                     CT053 Hypothetical Protein
                          (CT326)
                                     CT326 Similarity
                   0381
                          (CT047)
                                     CT047 Hypothetical Protein
                   0383
                   0387
                          (CT043)
                                     CT043 Hypothetical Protein
       40
                   0389
                          (CT041)
                                      CT041 Hypothetical Protein
                                      CT038 Hypothetical Protein
                   0393
                          (CT038)
                          (CT257)
                                     CT257 Hypothetical Protein
                   0395
                                      CT253 Hypothetical Protein
                          (CT253)
                   0399
                   0400
                          (CT254)
                                      CT254 Hypothetical Protein
        45
                   0401
                          (CT255)
                                      CT255 Hypothetical Protein
                                      CT105 Hypothetical Protein
                   0405
                          (CT105)
                                      CT102 Hypothetical Protein
                          (CT102)
                    0408
                                      CT260 Hypothetical Protein
                    0409
                          (CT260)
                          (CT262)
                                      CT262 Hypothetical Protein
                    0411
        50
                          (CT263)
                                      CT263 Hypothetical Protein
                    0412
                                      CT266 Hypothetical Protein
                          (CT266)
                    0415
                          (CT271)
                                      CT271 Hypothetical Protein
                    0420
                    0422
                          (CT273)
                                      CT273 Hypothetical Protein
                                      CT274 Hypothetical Protein
                    0423
                          (CT274)
        55
                                      CT276 Hypothetical Proteins
                          (CT276)
                    0425
                    0426
                          (CT277)
                                      CT277 Similarity
                                      CT283 Hypothetical Protein
                         (CT283)
                    0434
```

		0441	(CT007)	CT007 Hypothetical Protein
		0442	(CT006)	CT006 Hypothetical Protein
		0443	(CT005)	CT005 Hypothetical Protein
		0474	(CT365)	CT365 Hypothetical Protein
	· 5	0476	(CT865)	CT865 Hypothetical Protein
		0480	(CT383)	CT383 Hypothetical Protein
		0485	(CT382)	CT382.1 Hypothetical Protein
		0487	(CT384)	CT384 Hypothetical Protein
		0489	(CT386)	CT386 Hypothetical Protein
	10	0490	(CT387)	CT387 Hypothetical Protein
		0491	(CT389)	CT389 Hypothetical Protein
		0496	(CT391)	CT391 Hypothetical Protein
,		0497	(CT388)	CT388 Hypothetical Protein
		0506	(CT421)	CT421 Hypothetical Protein
	15			**
	13	0507	(CT421)	CT421.1 Hypothetical Protein
		0508	(CT421)	CT421.2 Hypothetical Protein
		0512	(CT425)	CT425 Hypothetical Protein
		0514	(CT427)	CT427 Hypothetical Protein
	20	0518	(CT429)	CT429 Hypothetical Protein
	20	0522	(CT433)	CT433 Hypothetical Protein
		0525	(CT398)	CT398 Hypothetical Protein
1		0533	(CT406)	CT406 Hypothetical Protein
Į,j		0537	(CT814)	CT814.1 Hypothetical Protein
Ü		0538	(CT814)	CT814 Hypothetical Protein
i ala	25	0554	(CT440)	CT440 Hypothetical Protein
i.		0559	(CT441)	CT441.1 Hypothetical Protein
I		0565	(CT449)	CT449 Hypothetical Protein
=======================================		0572	(CT456)	CT456 Hypothetical Protein
14		0582	(CT465)	CT465 Hypothetical Protein
	30	0583	(CT466)	CT466 Hypothetical Protein
		0588	(CT469)	CT469 Hypothetical Protein
H		0589	(CT470)	CT470 Hypothetical Protein
H		0590	(CT471)	CT471 Hypothetical Protein
١ <u>.</u>		0593	(CT474)	CT474 Hypothetical Protein
Ū	35	0595	(CT476)	- · ·
	55		(CT483)	CT476 Hypothetical Protein
		0601		CT483 Hypothetical Protein
		0602	(CT484)	CT484 Hypothetical Protein
	•	0606	(CT488)	CT488 Hypothetical Protein
	40	0609	(CT490)	CT490 Hypothetical Protein
	40	0622	(CT503)	CT503 Hypothetical Protein
		0623	(CT504)	CT504 Hypothetical Protein
		0648	(CT529)	CT529 Hypothetical Protein
	÷	0658	(CT538)	CT538 Hypothetical Protein
		0667	(CT546)	CT546 Hypothetical Protein
	45	0668	(CT547)	CT547 Hypothetical Protein
		0669	(CT548)	CT548 Hypothetical Protein
		0671	(CT550)	CT550 Hypothetical Protein
		0673	(CT552)	CT552 Hypothetical Protein
		0675	(CT696)	CT696 Hypothetical Protein
	50	0676	(CT695)	CT695 Similarity
		0681	(CT691)	CT691 Hypothetical Protein
		0687	(CT482)	CT482 Hypothetical Protein
		0688	-	• •
			(CT481)	CT481 Hypothetical Protein
	55	0700	(CT676)	CT676 Hypothetical Protein
))	0705	(CT671)	CT671 Hypothetical Protein
		0706	(CT670)	CT670 Hypothetical Protein
		0708	(CT668)	CT668 Hypothetical Protein

```
0709
                                              (CT667)
                                                                    CT667 Hypothetical Protein
                                 0710
                                               (CT666)
                                                                    CT666 Hypothetical Protein
                                                                    CT665 Hypothetical Protein
                                 0711
                                               (CT665)
                                                                    CT663 Hypothetical Protein
                                              (CT663)
                                 0713
              5
                                              (CT656)
                                                                    CT656 Hypothetical Protein
                                 0717
                                 0718
                                              (CT657)
                                                                    CT657 Hypothetical Protein
                                                                    CT659 Hypothetical Protein
                                 0720
                                              (CT659)
                                 0722
                                                                    CT654 Hypothetical Protein
                                              (CT654)
                                                                    CT652.1 Hypothetical Protein
                                 0725
                                              (CT652)
           10
                                 0726
                                              (CT620)
                                                                    CT620 Hypothetical Protein
                                 0727
                                              (CT619)
                                                                    CT619 Hypothetical Protein
                                 0739
                                              (CT638)
                                                                    CT368 Hypothetical Protein
                                 0742
                                              (CT635)
                                                                    CT635 Hypothetical Protein
                                 0746
                                              (CT632)
                                                                    CT632 Hypothetical Protein
           15
                                 0747
                                              (CT631)
                                                                    CT631 Hypothetical Protein
                                              (CT651)
                                                                    CT651 Hypothetical Protein
                                 0751
                                 0755
                                              (CT616)
                                                                    CT616 Hypothetical Protein
                                0760
                                              (CT611)
                                                                    CT611 Hypotheti:al Protein
                                 0761
                                              (CT610)
                                                                    CT610 Hypothetical Protein
der open door open to de service designations of the service designation of
         20
                                              (CT648)
                                0764
                                                                    CT648 Hypotheti:al Protein
                                0765
                                              (CT647)
                                                                    CT647 Hypotheti:al Protein
                                0766
                                              (CT646)
                                                                   CT646 Hypothetical Protein
                                                                   CT645 Hypotheti al Protein
                                0767
                                              (CT645)
                                0770
                                              (CT642)
                                                                   CT642 Hypotheti :al Protein
          25
                                0774
                                              (CT606)
                                                                   CT606.1 Hypothetical Protein
                                0776
                                              (CT605)
                                                                   CT605 Hypothetical Protein
                                0779
                                              (CT602)
                                                                   CT602 Hypothetical Protein
I
                                0783
                                              (CT598)
                                                                   CT598 Hypothetical Protein
ä
                                              (CT590)
                                0791
                                                                   CT590 Hypothetical Protein
H
          30
                                              (CT589)
                                0792
                                                                   CT589 Hypothetical Protein
H
                                0803
                                              (CT584)
                                                                   CT584 Hypothetical Protein
ļ÷
                                0807
                                             (CT580)
                                                                   CT580 Hypothetical Protein
H
                                0808
                                             (CT579)
                                                                   CT579 Hypothetical Protein
Ð
                                0809
                                              (CT578)
                                                                   CT578 Hypothetical Protein
₩ 35
                                0810
                                              (CT577)
                                                                   CT577 Hypothetical Protein
                                              (CT573)
                                                                   CT573 Hypothetical Protein
                                0818
                                             (CT569)
                                                                   CT569 Hypothetical Protein
                                0819
                                             (CT568)
                                                                   CT568 Hypothetical Protein
                                0820
                                             (CT567)
                                                                   CT567 Hypothetical Protein
          40
                                0821
                                             (CT566)
                                                                   CT566 Hypothetical Protein
                                0822
                                             (CT565)
                                                                   CT565 Hypothetical Protein
                                0827
                                             (CT560)
                                                                   CT560 Hypothetical Protein
                                0834
                                             (CT556)
                                                                   CT556 Hypothetical Protein
                                0840
                                             (CT700)
                                                                   CT700 Hypothetical Protein
          45
                                0842
                                             (CT702)
                                                                   CT702 Hypothetical Protein
                                0843
                                             (CT702)
                                                                   CT702 Hypothetical Protein
                                0852
                                             (CT711)
                                                                   CT711 Hypothetical Protein
                                0853
                                             (CT712)
                                                                   CT712 Hypothetical Protein
                                0857
                                             (CT716)
                                                                   CT716 Hypothetical Protein
          50
                                0859
                                             (CT718)
                                                                   CT718 Hypothetical Protein
                                0865
                                             (CT724)
                                                                   CT724 Hypothetical Protein
                                0869
                                             (CT728)
                                                                   CT728 Hypothetical Protein
                                             (CT733)
                                0874
                                                                   CT733 Hypothetical Protein
                                0875
                                             (CT734)
                                                                   CT734 Hypothetical Protein
         55
                                0884
                                             (CT741)
                                                                   CT741 Hypothetical Protein
                                0887
                                             (CT744)
                                                                   CHLTR Possible Phosphoprotein
                                0896
                                            (CT753)
                                                                  CT753 Hypothetical Protein
```

		0900	(C1/03)	C1703 H	ypomencar Protein		
		0908	(CT764)	CT764 H	ypothetical Protein		
		0912	(CT768)	CT768 H	ypothetical Protein		
		0925	(CT779)	CT779 H	ypothetical Protein		
	5	0938	(CT788)	CT788 H	ypothetical Protein		
		0939	(CT790)		ypothetical Protein		
		0943	(CT794)		Hypothetical Protein		
		0945	(CT795)		ypothetical Protein		
	10	0956	(CT805)		ypothetical Protein		•
	10	0960	(CT809)		ypothetical Protein		
		0989	(CT832)		ypothetical Protein	•	
		0994	(CT837)	CT837 H	ypothetical Protein		
		0995	(CT838)	CT838 H	ypothetical Protein		
		0996	(CT839)	CT839 H	ypothetical Protein		
	15	1002	(CT845)	CT845 H	ypothetical Protein		
		1003	(CT846)	CT846 H	ypothetical Protein		
		1004	(CT847)	CT847 H	ypothetical Protein		
•		1005	(CT848)	CT848 H	ypothetical Protein		
		1006	(CT849)	CT849 H	ypothetical Protein		.*
*sem	20	1007	(CT849)	CT849.1	Hypothetical Protein	•	
ii		1008	(CT850)	CT850 H	ypothetical Protein		J
111		1010	(CT852)		ypothetical Protein		ş.
-		1011	(CT853)		ypothetical Protein		ž.
		1015	(CT857)		ypothetical Protein		1
gerij (m.) il di semp genje eli busis der dete maj dang ester	25	1016	(CT858)		ypothetical Protein		*
\$=\$	23	1019	(CT860)				÷
II		1020			ypothetical Protein		d S
			(CT861)		ypothetical Protein		
ä		1022	(CT863)		ypothetical Protein		
	30	1032	(CT373)		ypothetical Protein		
ļ	50	1033	(CT372)		ypothetical Protein		
E.s.		1034	(CT37i)		ypothetical Protein		
		1057	(CT356)		ypothetical Protein		
		1058	(CT355)		ypothetical Protein		
	25	1061	(CT330)		ypothetical Protein		
ħ.J	35	1073	(CT371)	CT371 H ₂	ypothetical Protein		
					Coding Control of Control		
		0486	•		Coding Genes Not in C. trachomatis Hypothetical Proline Permease		
		0279					
	40	0505			Possible ABC Transporter Permease Protein		
	70				3-Methyladenine DNA Glycosylase		
		0193		argR	Similarity to Arginine Repressor		
		1041		bioA	Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase		4
		1044		bioB	Biotin Synthase		
	15	1042		bioD	Dethiobiotin synthetase		
	45	0585			Similarity to Cps IncA_2		
		0562			CHLPS 43 kDa Protein Homolog_1		
		0927			CHLPS 43 kDa Protein Homolog_2		
		0928			CHLPS 43 kDa Protein Homolog_3		
	50	0929			CHLPS 43 kDa Protein Homolog_4		
	50	1045			Conserved Hypothetical Membrane Protein		
		0251			Conserved Hypothetical Protein		
		0278			Conserved Outer Membrane Lipoprotein Protein		
		0907			CutA-like Periplasmic Divalent Cation Tolerance Protein		
		0171		guaA	GMP Synthase		
	55	0172		guaB	Inosine 5'-Monophosphase Dehydrogenase		
		0608			Uridine 5'-Monophosphate Synthase		
		0735			Undine Kinase		

0906 (CT763) CT763 Hypothetical Protein

		0980		Similar to Sac	charomyces cerevisiae	52.9KDa Prote	ein				
		0232			'-Methylthioadenosine				•		
		1046		Tryptophan H							
		0477	vaeV Bs		•						
	5				pothetical Protein						
	,	0048	yqfF-Bs	•	pothetical IM Protein						
		0587	yvyD_Bs	Conserved Hy	pothetical Protein						
		0143	yxjG_Bs_	1 Conserved	Hypothetical Protein						
		0448	yxiG_Bs_	2 Conserved	Hypothetical Protein						
		0006	0180	0440	0977						
	10	0007	0181	0455	0978					,	
		0008	0190 -	0456	1018					,	
		0009	0203	0457	1023						
		0010	0204	0458	1027						
		0011	0205	0459	1029						
	15	0012									
	15		0209	0460	1040						
		0028	0210	0461	1051						
		0029	0211	0462	1052						
		0034	0212	0463	1053						
		0041	0213	0464	1054						
O	20	0042	0214	0465	1055						
ñ		0043	0215	0472	1056						
Trade gran		0044	0216	0473	1064						
1.1		0045	0218	0481	1065						
14		0046	0220	0483	1066						
the first the many time with the party lines.	25	0047	0221	0492	1070						
j-li		0049	0222	0493							•
(I)					1071			*			
Į.		0050	0223	0494	1072						3
8		0051	0224	0498							
La	20	0063	0225	0499		-	*				
4	30	0064	0226	0516							
E :		0066	0233	0517							
ļ.		0067	0240	0523							
ļeb		0069	0241	0524							
		0070	0242	0553							
Ð	35	0099	0243	0574						•	
		0124	0266	0600							
		0125	0267	0656							
		0126	0268	0664							
		0130	0277	0677							
	40	0131	0283	0678							
		0132	0284	0685							
		0132	0284								
				0686							
		0146	0287	0724			•				
	45	0147		0731							
'	43	0155		0745							
		0156		0753							
		0157	0355	0794							
		0158	0356	0795							
		0159	-0357	0796					:		
	50	0162	0358	0797							
		0163	0365	0798							
		0164		0799							
		0165		0829							
		0166		0830							
	55	0167		0831		•					
		0168	0372	1880							

0170	0376	0913
0173	0391	0914
0174	0398	0930
0175	0404	0944
0177	0431	0964
0178	0432	0975
0170	0420	0076

As A2 Table 3

Chlamydia pneumoniae Genome Encoded Pr

CPn_0001 330 4
CT001 hypothetical protein
KRLKDELKYTSLRKKANLCKIIRGLSSLIVILCALNVGLIGITHNKLNIIAKLCGGVSTP
ATQITYIIIGIAGVICLLSFCPFCSKKSRHSHGDSCSSGGCHSHHSDKN

CPn_0002 570 875
sare into repart to the foreign of labority and enterprise of labority and enterprise of the foreign of the property of the p

CPn_0003 889 2370
gata-Glu trna Gln Amidotransferae
KIMYRYSALELAKAVTLGELTATGVTQHFFHRIEEAEGQVGAFISLCKEQALEQAELIDK
KRSRGEPLGKLAGUPVGIKDNIHVTGLKTTCASRVLENYQPFFDATVVERIKKEDGIILG
KLMNDEPAMGSTTLYSAFHPTHNEWDLSRVPGGSSGSGAAAVSAFCEVALGSDTGGSIR
OPAAFCGVVGFKESYGAVSRYGLVAFASSLDQIGPLANTVEDVALMDDYFSGRDFKDATS
REFFRDSFMSKLSTEVPKVIGVPRTFLBGLRDDIRENFFSSLAIFEGEGTHLUTVELDIL
SHAVSIYYILASAEARINLARFGGVKYGYRSPOAHTISQLVDLSRGBGFGKEVMRRILLG
NYVLSAERQNVYYKKATAVRAKIVKAFRTAFEKCEILAMPVCSSPAFEIGEILDPVTLYL
QDIYTVAMMLAYLPAIAVPSGFSKEGLPLGLQIIGQQGDQOVCQVGYSFQEHAQIKQLF
SKRYAKSVVLGGOS

CPn_0004 2334 3833
gatB-(Petl12) Glu trNA Gln Amidotransferase (B Subunit)
EICQKCCSRRSIMSAVYADMESVIGLEVHVELNTASKLFSSALNRFGDEPNTNISTVCTG
LPGSLEVLNQSAVEKAVLFGCAVEGEISLLSRFDRKSYFYPDSPRNFQITOFFEHPITRGG
RIKAIVQGERTYFELAQTHIEDDAGHLKHFGEFAGVTYNRAGVPLIEIVSKPCMFCPEDA
AVAYATSLVSLLDYIGISDCNMEEGSIRFDVNVSVRPKGSFELNKVEINDMNSFARMAQA
LEAEKQRQIDEYLNQPNKDPKLVIPAATYRWDPEKKKTVLMRLKESAEDYKYFPEPDLPT
LQLTESYIERIRKTLPELPYDKYHRYJQEYGLSEDIASILISKNIATFEVACKDCKNF
RSLSNWATVEFGGRCKTLGVKLPSSGIFPEGVAQLVNAIDQGVITGKIAKEIADLMMESP
GKNPEEILKEKPELLPMSDEGELQKIIAEVVLANPESIVDYKNGKTKALGFLVGQIMKRT
AGKAPPKRVNELLLLELDKG

CPN_0005 4097 6892

pmp_1-Polymorphic Outer Membrane Protein
SDIHFDLGTKMRFSLCGFPLVFSFTLLSVFDTSLSATTISLTPEDSFHGDSQNAERSYNV
OAGDVYSLTGDVSISNYDNSALMKACFNVTSGSVTFACNHHGLYFNNISSGTTKEGAVLC
CQDPQATARFSGFSTLSFIGSPGDIKEQGCLVSKNALMLLLNYVVAFFEQNGSKTKGGAIS
GARUTIVGNYDSVSFYQNAATFGGAIHSSGPLOIAVNQAEIRFAQNTAKNGSGGALVSDG
DIDEDQNAYVLFRENEALTTAIGKGGAVCLEFTSGSSTPVPIVTFSDNKQLVFERNHSIM
SCGATYARKLSISSGGPTLFINNISYANSQNLGGAIAIDTGGEISLSAEKGTITFGCNRT
SCGETSLANDPRDFKSTIPQNVNLSAGYLVIKEGAEVTVSKFTQSFSSHLVLLOTKLIAS
SGETSLANDPRDFKSTIPQNVNLSAGYLVIKEGAEVTVSKFTQSFSSHLVLLOTKLIAS
KEDIAITGLAIDIDSLSSSSTAAVIKANTANKQISVTDSIELISPTGNAYEDLRMRNSQT
FPLESLEPGAGGSVTVTAGGFLFVSFHYGFQCNMKLAMTGTGNKVGEFFMDKINYKPRPE
KEČMLVPNILMGNAVDVRSLMQVGETHASSLCTDRGLMIDGIGNFHVSASEDNIRVRHN
SGGTVLSVNNEITPKHYTSMAFSQLFSRDKDYAVSNNEYRMYLGSYLYQYTTSLGNIFRY
ASRMPNVNVGILSRRFLQNPLMIFHFLCAYGHATMDMCTDYANFPMVNSWRNNCMAIEC
GSMPLLVFENGELFGGAIFFMKADLVVAYYGGFKRTTADGRFSRNSGLTSISVPLGIFF
KLALSQDVLYDFSFSYIPDIFRKDPSCRAAVISGDSWLVPAAHVSRHAFVGSGTGRYH
FNDYTELLCRGSIECRPHARNYNINCGSKFRF

CPT_0006 7299 7141 No robust homolog present in Genebank/EMBL as of 11/7/98 KQEQEPLRSALLERLSEWLVLLGVPSPETTRSTPEKDANQLPKDSRNRTLESL

TO TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL T

CPn_0008 10780 11685
NO robust homolog present in Genebank/EMBL as of 11/7/98
CKYSYLLNYPPPRRSLCVSCSKLRSLSITLLVLGVLLLTLGIFGLTAGISFGAGLGFSA
LGGVLVISGLLFLLVRREVETVRSEEIPRGVSVTPSEEPALEWAGKEPETKKILDRLEKE
LDQLDTYIQEVFACLERLKDPKYEDRGLLTEAKEKLRVFDVVEKDMMSEFLDIQRVLNEE
AYYVEHCQDPLENIAYEIFSSOELRDYYCAGVCGYLPSGDAADRLKRSVKEVMDRFMRV
TWKSWEASVMLDHSYGVARELFKKAVGVLEESVYKILFKSTRDAFYECEKAKIQRDGFFK

CPn_0009 11689 13119
No robust homolog present in Genebany/EMBL as of 11/7/98
-fitsahaborerding:medikot-lewygehoctdiffyarks:cmmddbyadkfilrekeekmerheldehatmyrkaschayakakaafekersnebbekykdvekwlskglaefragsbra
karerlreldylyfeysveervlerortkkynleny/kadiekkyhhicyregehymkevenk
Paeyremiekylsaefyskclorlebolletwskyltkaebyfemkfdateklonkylsd
ythrietilsoehierekhololyfaythygenehyfbateklonkylsd
ythrietilsoehierekhololyfaythygenehyfbatekorymiskemlakyepo
ythrietisoehierekhololyfaythygenehyfbatekorymiskemlakyepo
liffingellaviaelftoarlolyatypymefyloyhrikrekynsmakteryretro
afysymkebliaedtilkeedywlipdbwlyrderknorriiskaaooevkgf

Cim_0010 13324 19325 No robust homolog present in Genebank/EMBC as of 11/7/98 CKYFYLRSYPPPPHSVIS PROBLEM RELEVED LLISAAIS FGLGIGLSALGVLMISGLL REIPTVRPEEIPSCVSLAPSEEPSALQAAGKTLAQL REIPTVRPEEIPSCVSLAPSEEPSALQAAGKTLAQL PKELDOLDTDIGSVFACLRKLADSKYSSRSFLNDAKKELRVFDFV/EDTLSSEIFELRQIV AQESHDLNFLINGGRSLMMTASSESLDLFHVSKRJGYLPSGDVRGEGLKKSAKRIVARLM SLHCEIHKVAVAFDRNSYAMAEKAFAKALGALEESVYRSLTQSYRDKFLESERAKIPWNG HITWLRDDAKSGCAEKKLGMPRNVGRNLGKQSFG

CPn_0010.1 14268 15746

CPn_0010.1 14268 15746

CPn_modific with the company of t

CPn_0011 1587 16614

gatB-(Pet112) Glu trix Gln Amidotransferase (B Subunit)

fwysimtaapailhvsprppfetrkrviekdskskalgitlluvgliuvggaivlsgvis

GlsalivcGlgistislgvufrvLglilllkrkeltleQlfakQlaeftapelkelemyi

OSTEKSLEKIEGSRYSDOFFLNRATQKILDLESSLSSITSEFRDLRQLFDEEKIELEMYI

OSTEKSLEKIEGSRYSDOFFLNRATQKILDLESSLSSITSEFRDLRQLFDEEKIELLSGE

RLLEFTAANLFKQGRDVYLNLGNLADIRAYMGPNNYKVAMVIEKAKAVVHEFIVLTTMAR

FLEFFE

CPn_0012 16596 18212

gatB-(Pet112) flu trna Gln amidotransferase (B Subunit)

girvfflknkyglikgmygehlrilerilynsvyksyaprifsyertmyhhdtplipwee

kekcaeaekapiegokilldygksifwinendeinilndpwswgimtvarrkvygevdds

erwhkvligkjeddyeklleesskesteankklisdlydriedatkfflikgeevetr

vkdlaaryggtydekgoteakkvyeleasletfildsieselvocledgdiywkegovykd

lattoeleeddeakkeaeaaedlrsinerikkskthildrakwhienaedsitwhtsgiemk

dmkarlkilkeditsvleeideietclsleeiptlttrelltrkykkfkicsetlikmts

vfenniyvgeyevolonigfklogisgapskkoddpfanleedvalgkkriketionfeig

gfnymkedfkaaakdlyirstaegkmnfdvpcmelfrryheevnkplleimycadsyrd

akkkilcslridekellokeikkeefyokkogrhadrsrhttygklriaeelalelkkki

CPI_0013 18509 21106
pmp_J-Polymorphic Outer Membrane Protein
RRDPLAFFIYLLYWKESPLREKKVMKIPLRFILLISLVPTLSMSNLLGAATTEELSASNS
FROMTAFFIYLLYWKESPLREKKVMKIPLRFILLISLVPTLSMSNLLGAATTEELSASNS
FROMTSTTISFSSKTSSATDGTNYVFKDSVVIENVPKTGETQSTSCFKNDAAAGDLNFLGG
GF\$FFFSNIDATTASGAAIGSEAANKTVTLSGFSALSFLKSPRSTVTNSLGAINVKGNLS
LENDRVLIQDNFSTGATGGGAINCAGSLKIANNKSLSFLGNSSSTRGGAIHTKNLTLSSGG
ETLFQGNTAPTAAGKGGAIAIADSCTLSISGDSGDIIFEGNTIGATGTVSHSAIDLGTSA
KITALRAAQGHTIYFYDETVTGSTSVADALNINSPDTGDNKEYTGTIVFSGEKLTEAEA
KEDRINTSKLLONVAFKNSTVVLKGDVVLSANGFSQDANSKLIMDLGTSLVANTESIELT
NLEINIDSLRNGKKIKLSAATAQKDIRIDRPVVLAISDESFYQNGFINEDHSYDGILELD
AGKDIVISADSRSIDAVQSPYGYQGKWTINWSTDDKKATVSNAKQSFNPTAEQEAPLVM
LLMGSFIDVRSFQNFIELGTBCAPYEKRFWVAGISNVLHRSGRENQRKFRHVSGGAVVGA
SYRMGGDTLSLGFAQLFARDKDYFMYTNFAKTYAGSLRLQHDASLYSVVSILLGEGGLR
ETLLPYVSKTLPCSFYGOLSYGHTDHRWKTESLPPPPPTLSTDHTSWGGYVWAGGLGTRV
AVENTSGRGFFQEYTPFVKVQAVYARQDSFVELGAISRDFSDSHLYNLAIPLGIKLEKRF
ABQYYHVVAMYSPDVCRSNPKCTTTLLSNQGSWKTKGSNLARQAGIVQASGFRSLGAAAE
LFGNFGFEWRGSSRSYNVDAGSKIKF

CPn_0014 21365 21922
pmp_3-Polymorphic Outer Membrane Protein
IONOSITYFTMKSSFREYPSTFAIFPLSMIATETV/LDSSASFDGNKNGNFSVRESQEDAG
TTYLFKGNVTLENIPGTGTAITKSCFNNTKGDLTFTGNGNSLLFGTVDAGTVAGAAVNSS
VVDKSTTFIGFSSLSFIASPGSSITTGKGAVSCSTGSLSLTKMSVCSSAKTFQRIMAVLS
POKLFH

CPn_0015 21835 24174
pmp_3-PMP_3 (frame-shift wich 0014)
LEFDKNVSLLFSKNFSTDNGGAITAKTLSLTGTTMSALFSENTSSKKGGAIOTSDALTIT
GNGGEVSFSDNTSSDSGAAIFTEASVILISNNAKVSFIDNKVTGASSSTTGDMSGGAICAY
KTSTDTKVTLTGNQMLLFSNNTSTTAGGAIYVKKLELASGGLTLFSRNSVNGGTAPKGGA
IA IEDSGELSLSADSGDIVFIGNTVTSTTPGTNRSSIDLGTSAKMTALRSAAGRAIVFYD
PITTGSSTTVTDVLKVNETPADSALQYTGNIIFTGEKLSETEAADSKNITSKLLQPVTD
GGTLSLKHGVTLQTQAFTQQADSRLEMDVGTTLEPADTSTINNLVINISSIDGAKKAKIE
TKATSKNLTLSGTITLLDPTGTFYENHSLRNPQSVDILELKASGTVTSTAVTPDPIMGEK
FHYGYQGTMGPIWGTGASTTATFMTKTTGYIPNPERIGSLVPNSLUMAPIDISSLHAFT
ETANEGLQGDRAFWCAGLSNFFHKDSTKTRGFRHLSGGYVIGGNLHTCSDKILSAAFCQ
LFGRDROFFVAKNGTVYGGTLYYQHNETYISLPCKLRPCSLSYVPTEIPVLFSGNLSYT
HTDNDLKTKYTTYPTVKGSWGNDSFALEFGGRAPICLDESALFECYMPFMKLOFVVAHQE
GFKEQGTEAREFGSSRLVNLALPIGIRFDKESDCQDATVNLTLGYTVDLVRSNPDCTTTL
RISGDSWKTFGTNLARQALVLRAGNHFCFNSNFEAFSQFSFELRGSSRNYNVDLGAKYQF

CPD_0016 24383 25188
pmp_4-Polymorphic Outer Membrane Protein
RSDFALKROCHMRSSFSLLLISSSLAFFLLMSVSADAADLTLGSRDSYNGDTSTTEFTPK
AATSDASGTTY/ILDGDVSIS-SAGKGTSLTTSCFSNTAGNLTFLGNFSSHFONLISSTVA
GVVVSNTAASGITKFSCFSTLRMLAPPTTTKGATKILVEPSIGNLDLNENASSEN
GAINTKTLSLTGSTRFVAFLSNSSSQCGATYASGDSVISENAGILSFGNNSATTSGGAT
SAEGNLVISNNQNIFFDGCKATTIFGATLCNKAGANPDFILTLSGESHHFLNHTAGNSG
GAIYTKKLVISSGRGGVLFSNNKANATPKGGATATLDSGETSISADLGNITFEGNTTST
TGGPAGVTRNAIDLASNAKFLNLPATPGGIKVIFYDPTTSSKATDKLSLHKADAGSGNTYS
TYTYFGGEKLSEELKKPDNLKGTFTDAFELAAGALVLKDGVTVVANTTTQVEGSKVYMD
GEGTEFFEASAHSVELNGLAINIDGLDGTIKALKATAASKDVALKAGTHVVANTTTQVEGSKVYMD
LDGOGVEFPLIELSAGATMTTTETIPTTHYTYGYKNMMNGLGTRETNCKNKKCYLN

CPO_0017

CPO_00

CPn_0018 27513 27003

pmp_5-Polymorphic Outer Membrane Protein
EYNMKTSVSMLLALLCSGASSIVLHAATTPLHPEDGFIGEGNTNTFSPKSTTDAAGTTYS
EYNMKTSVSMLLALLCSGASSIVLHAATTPLHPEDGFIGEGNTNTFSPKSTTDAAGTTYS
ETGEVLYIDPCKGGSITCTCFVETAGDLTFLGNGNTLKFLSVDAGANIAVAHVQGSKNLS
FTDFLSLVITESPKSAVTTGKGSLVSLGAVQLQDINTLVLTSNASVEDGGVIKNSCLIQ
GIKNISAIFGONTSSKAKGAISTTOLTIENNLTTLKFNENKAVTSGGALDLGASFFTAN
HELLESJUETTSGARAID BALDE DE STEPTATTCLLGEDETTMEDVAALCTTTTGITCSD
GILVFEGNQVTTTAPNATTKRNVTHLESTAKWTGLAASGNAIYFYDPITTMDTGASDN
LRINEVSANOKLSGSIVFSGERLSTAEAIAENLTSRINQPVTLVEGSLVLKQGVTLITQG
FSQEPESTLLLDLGTSL

CPn_0019 29007 30356
pmp_5-pmp_5 (frame-shift with 0018)
ASTEDIVITHLSINADTIYGKMPINIVASAANKNITLTGTLALVNADGAFYENHTLQDSQ
DYSFVKLSPGAGGTIITODASCRPLEVAPSRPHYGYQGHWNVQVIPGTGTQPSQANLEWV
RTGYLPNPERQGSLVPNSLWGSFVLOGRAIQEIMVNSSOILCQERGVWGAGIANFLHRDKI
NEHGYRHSGVGYLVGVGTHAFSDATINAAFCQLFSRDKDYVVSKHHGTSYSGVVFLEDTL
EFRSPCGFYTDSSSEACCNQVVTIDMQLSYSHRNNDMKTKYTTYPEAQGSWANDVFGLEF
GATTYYYPNSTFLFDYYSFFLRQCTYAHQEDFKETGGEVRHFTSGDLFNLAVPIGVKFE
RFSDCKRGSVEITLAVYPDVIRKDPKSTATLASGATWSTHGNNLSRQGLQLRLGNHCLIN
PGIEVFSHGAIELRGSSRNYNINLGGKYRF

CPn_0020 32717 30603

Predicted OMP [leader (14) peptide: outer membrane]

KLWSNPNLRLMKRCFLFLASFVLMGSSADALTHOEAVKKKNSYLSHFKSVSGIVTIEDGV

LNIHNNLRIQANKVYVENTVCQSLKLVAHGNYMVNYRAKTLVCDYLEYYEDTDSCLLTNG

RFAMYPWFLGGSMITLTPETTVIRKGYISTSEOPKKDLCLSGDYLEYSSDSLLSIGKTTL

RVCRIPILFLPPFSIMPMEIPKPPINFRGGTGGFLGSYLGMSYSPISRKHFSSTFFLDSF

FKHGVGMGFNLHCSGKQVPENVFNKKSYYAHRLAIDMAEAHDRYRLHGDFCFTHKHVNFS

GEYHLSDSWETVADIF PNNFMLKNTGPTRVDCTWNDNYFEGYLTSSVKVNSFONANGELP

YLTLRQYPISIYNTGVYLENIVECGYLNFAFSDHIVGENFSSLRLAARPKLHKTYPLPIG

TLSSTLGSSLIYYSDVPBISSRHSQLSAKLQLDYRFLLHKSYIQRRHIIEPFVTFITETR

FLAKNEDPYIFSIODAFHSLNLLKAGIDTSVLSKTNPRFPRIHAKLWTTHILSNTESKPT

FPKTACELSLPFGKNTVSLDAEWIWKKHCWDHMNIRWEWIGNDNVAMTLESLHRSKYSL

IKCDRENFILDVSRPIDQLLDSPLSDHRNLLLGKLFVRPPPCWNYRLSLRYGWHRQDTPN

YLEYOMULGTRIFEHWQLYGVYEREADSRFFFFLKLDKPKKPPF

CPH-0021 34470 32707
Predicted OMP [leader (19) peptide]
CSRSPYPNIEILARGVEHRSMGLFHLTLFGLLLCSLPISLVAKFPESVGHKILYISTOST
ONLATYLEALDAYGHDFFVLRKIGEDYLKQSIHSSDPOTRKSTIIGAGLAGSSEALDV
LSQAMETADPLOQLLVLSAVSGHLGKTSDDLLFKALASPYPVIRLEAAYRLANLKNTKVI
DHLHSFIHKLPEEICCLSAAIFLRLETEESDAYIRDLLAAKKSAIRSATALQIGEYQKR
FLEPHERRLLTSASPODQEAILYALGKLKDGOGSYNIKKQLQKPDVDVTLAAQAQLIALGK
EEDALPVIKKQALEERPRALYALRHLPSEIGIPIALPIFLKTKNSEAKLNVALALLELGC
DTPELLEYITERLVQPHYNETLALSFSKGRTLONKRVNIIVPODPOERERLLSTTRGLE
EQILITERLFRENEAYLEPIYKLLASQKTQLATTAISFLSHTSKQEALDLLFQAKLGEP
IIRAYADLAINLTKDPEKKRSLHDYAKKLIQETLLFVDTENQRPHPSMPYLRYQVYPES
RTKYALDILETLATSKSSEDIRLLIQLMTEGDAKNFPVLAGLLIKIVE

CP0 0022 35042 34395 maf

TIEQVISNCCNVSNTRSFYSMSLPLVLGSSSPRRKFILEKFRVPFTVIPSNFDESKY GDPLAYTOELAAQKAYAVSELHSPCDCIILTGDTIVSYDGRIFTKRODKADAIQMLK NOGTFDVYTSIAVLHKGKLLTGSETSGJISLTMIPDHRIESYIDTVGTLNNCGAYDVCHC ILKAVHGCVYNVQGLPIQTLKYLLEELNIDLWDYSI

CPE_0023 36657 35014
yjjK/alr-abc Transporter Protein Atpase
ENAKLLYSKOHTVMLSAMSIVLDKIGKSLGTRILFDDVSVVFNPGNCYGLTGPNSAGKS
TLLKIIMGMIEPTRGSISLPKKVGILRONIDSFHDTTVLDCVIMGNTRLWEALDJRONLY
LOEFTDALGMELGEIEBIGEENCYRADSEAEBLLTGGIPNEMPFKKMAH PJOLOFRV
LLCOALFGHPEALLLDEPTNHLDLYSINWLGNFLKDYEGTVIVVSHDRHFLNTATTHIAD
IDVDTIIIYFGNYDDWYENKTASRODEKADIRSKEKKISQLKEFVAKRGASSASOVOSR
LREĪKKLOPOELKKSNIQRPYIRFPLSDKSSGKVVLSLEAITKDVGDHQVIHPFSLEIYQ
GDKLGIIGNNGLGKTTLMKLLGAVEAPSSGSIKLGHQAICSYFPONHSDVLADCOETLF
EWLRNKRTGINDGLIRSVLGKHLFGGDDAFKGIQALSGGETRALLMAGMWIENNVLILD
EANNHLDLESVSALSWAINDYKGTAIFVSHDRGLIQDCATKLLIFDKDKIFFDGTWVDY

CPL_0024 37605 36661
xerc-Integrase/recombinase
REVMIASIYSFLDYLKMVKSASPHTLRNYCLDLNGLKIFLEERGNLAPSSPLQLATEKRK
VSELPFSLFTKEHVRHYIAKLIENGKAKRTIKRCLSSIKSFAHYCVICOKILLENPAETIH
GPRLPKELPSPMTYAQVEVLMATPDISKYHGLRDRCLMELFYSSGIKISETVAVNKOPFD
LSTHLIRIRGKGKKERIIPVTSNAIQWIQIYLNHPDRKRLEKDPQAIFLNRFGRRISTRS
IDRSFQEYLRRSGLSCHITPHTIRHTIATHWLESGMDLKTIQALLGHSSLETTTVYTQVS
VKLKKOTHOEAHPHA

CPn_0025 38610 37684
elaC/atsA-Sulphohydrolase/Glycosulfatase
LLMSSRELITIACCSQOPTRTRNQGAYLFRWNGEGLLFDPCGTQROFIFANIAPTTVNR
IFVSHFHGDHCLGLGSMLMRLNLDKVSHPIHCYYPASGKKYFDRLRYGTIYHETIQVVEH
PISEEGIYEDFGSFRIEAQRLDHOVPTLGWRITEPDTIKHLPKELESRGIRGLIIQDLIR
DOELSIY/JSTVYLSDVSYVRKGDSIAIIADTLPCQAAIDLANNSCMMLCESTYLEQHRHL
AESHFHMTAKQAATLAKRAATQKLILTHFSARYLNLDDTYKEASAVFPNVSVAQEYRSYP
FPKNPLLNK

CPH_0026 39637 38762
CT 45 hypothetical protein
cnfamehilipsemenytsyfikpopikoaapskelydichtayelletoluv/yvlvgame
cmftenytsyfikpopikoaapskelydichtayeleletoluv/yvlvgame
cmftenytsteleesiaelveletippeetimmisykrykelapkoasesopykkoaskko
tolshhibippimfiktolekounyytnikpiksgefeihiddenisproskredetesode
coleelipkaagikahtaytikesktetteesikkkyktkhelhritosilikroapkpmvpsk
khkuvlikktyplpleduseogsgebesobssyptvorkalepmytkoptop

CPn_0027 42252 8977 34779
Ton Lon ATP dependent Protesse

PSIRTIVDSTTNSDSPILDI

CKLLDESEESEDQSTERLLPSELFILPLNKRPFF
PCMAAPILIESGPYYEVLK.

OKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPIECSSAQVLLSIEERIRTTETIKDKYLKARVSYHADNKELTEELKAYSISIVSVIKD
LLKLNPLFKEELOIFLCHSDFTEFCKLADFSVALTTATREELQEVLETTNMHORIDKALI
LLKKELDLSRLQSINQKIPATITKSGKEFFLKECLKTIKKEKELGLEKEDRAADIDIEKSE
LRKRHVPDYAMEVIQDEIEKCOTLETSSAEYTVCRNYLDWLTIIPWGIQSKEYHDLKKAE
IVILKDHYGLDEIKOR ILBLISVGKLSKGIKGSIICLVCPPCVGKTSIGRSIAKVLHRKF
PRFSVGGMRDEAEIKGHRFTYIGAMPGKMVOALKOSQANHPYHIDEVDKIGASYHADPA
AALLEILDBERNKOFELFUL BANCH, DEBALEFMEINNYAREAGVETLAFINIKKTLRKVA
LKIVQNQEKPKSKKITFKIOSKNLAFTYLDKPIFTSGDRFYESTPVGVATGLAWTSLGGATSLGAV
LIESVQNSSLKTDHHLTGQAGEVMKESSGIAWTYLHSALHRYAPGYTFFPKSQVHIHIPE
GATPKDGPSAGITMYTSLLSILLETPVVNNLGMTGEITLTGRVLGVGGIREKLIAARRSR
LNILIFPEDNRRDYEELPAYLKTGLKIHFVSHYDDVLKVAFPKLK

CPn_0028 43328 42543
No robust homolog present in Genebank/EMBL as of 11/7/98
RMFLOFFHPIVYSDQLSFLPYLGKSSGIIEKCSNIVEHYLHLGGDTSVIITGVSGATFL
SYDHALPISKSEKIIKILSYTLILELILALFIKIVLRIILFFKYRGLILDVKKEDLKKTL
TPDQENLSLPYPSPTTLKKIHALHILVRSGKTYNELIQEGFSFTKTTDLGQAFSPKQDIG
FSYNSLLPNFYFHSLVSVPNISGEERALNYHKEQQEEMAVKLKTMQACSFVFRSLHLPSM
QTKDKKAGFGLLTFFPWKIYPL

CPn_0029 43839 43390
No robust homolog present in Genebank/EMBL as of 11/7/98
SIKNERUMENIYCFNLFRYIRFFAALNIRMNDGLRFCYSVILLRPMLLDSSLLRKGGGELL
KKFOIKLRTTSIKSSLISLRQQLGKREATQSDILYGTSRFQYLNSFEIEDPRIPPTMAAQ
LQEITYSRSVMELKIKFYVYLNSERNKTKP

CPn_0030 43840 44529
gcp/O-Sialoglycoprotein Endopeptidase
LKQVCWYSLFFYIKNRRMYFYKYVIIDTSGYYPFLACVDNQQVLEHWSLPVGPDLGIVLE
FLFKSKNLSFQGVAVALGPGNFSATRIGISFAQGLAMAKNVPLLGYSSLGYYLLSKDEKK
AFMLPLGKRGGVLTLSSEIPEEGLNEKRRGVGPGALLSYEEASDYCVAHGYYHVISPNPQ
LFASSFSDKITVEEVAPSVDQIRRNVISQFMFVEYDKQLSPDYRSYSCIF

CPn_0031 44708 44884 rs21-S21 Ribosomal Protein CMPSVKVRVGEPVDRALRILKKKIDKEGILKAAKSHRFYDKPSVKKRAKSKAAAKYRSR

CPn_0032 44881 46098
dnaJ-Heat Shock Protein J
SLIGNVVFVGSVSGMDYYSILGISKTASAEEIKKAYRKLAVKYHPDKNPGDAAAEKRFKE
VSEAYEVLSDPOKRDSYDRFGKDGPFAGAGGFGGAGGMGNMEDALRTFMGAFGGEFGGGS
FFDGLFGGLGEAFGMRSDPAGARQGASKKVHINLTFEEAAHGVEKELVVSGYKSCETCSG
GAVNPGGIKSCERCKGSGOVVOSRGFFSMASTCPECGGEGRIITDPGSSCRGGGRVKDK
RSVHVHIPAGVDSGMRLKMEGYGDAGONGAPSGDLYVFIDVESHPVFERRGDDLILELPI
GFVDAALGMKKEIPTLLKTEGSCRLTVPEGIOSCTILKVRNQGFPNVHGKGRGDLLVRIS
VETPQNLSEEQKELLRTFASTEKAENFFKKRSFLDKIKGFFSDFTV

CPn_0033 46129 48171
pdhA&B/odbaLodbB-(pyruvate) Oxoisovalerate Dehydrogenase Alpha
& Beta Fusion
ERSMJVVQNQVISSIRDVLKLVWELRFAEHKMLLLSRQSGSGTFQLSCAGHELAGVLAG
KSLIPGKDWSFPYYRDQGFPIGLGCDLSEIFASFLARTTPNHSSARMMPYHYSHKKLRIC
CQSSVVGTOFLQAAGRAWAVKHSSADEVVYVSGGDGATSQGEFHEMLNFVALHQLPLITV
IQNNHWAISVPFEDQCGADLASLGRCHQGLAVVEVDGGNYTSLTETFSHAVDQARQHSVP
ALILIDVVRLSSHNSDNDEKYRSALDLKLSMDKDPLILLEKEAINVFGLSPFEIEEIKA
EAQDEVKRSCEIAEALPFPSKGTSTENEVFSPYTETLIDYENSESAQNLRNSEFKVMRDAI
SEALVEDMTRDSQIVFGEDVAGDKGGVFGVTRNLTEKFGPQRCFNSPLAEATIIGTAIG
MALDGIHKPVVELQFADYIWPGINQLFSEASSIYYRSAGEWEVPLVIRAPSGGYIQGGPY
SHOVLPFGKAAIVHPGKDLTIVSWGMPLVLSLEVAQELASRGISIEVIDLRTMVPCDFA
TVLKSLEKTGRLLVIHEASEFCGFGSELVATMSEQGYAYLDAPIRRLGGLHAPVPYSKVL
ENEVLPHKESILQAAKSLAEF

CPn_0034 49496 48210
CT345 hypothetical protein
nyfllpytrogilmæistpslpdssivsqktppvpdpdsspdhiptiptqapfkpqkk
etpssivnaiafallaflsclggvfaiclgcsleitmplfiltavfiaftllyfihylek
pkipeplptpppsptlraptltpeipapapgiplpptlpkvdrtkltonpdihypstydp
kacfsllkqlfsldpetrpedrkysnklasillrskeksgfrfhcfkghfshdkllnkks
gavvisshssmdfsttlgrafavttclqrscwekiknniptpekhlpigscvsgpmdvee
gaqlytshlivinpptletlikekmrraitlkdfsmkeafthlvlaylqcfdiciehnle
svqlevfglnnlsadqeefttwescchlallesvrillaskeeyalsnvsvnsisqvplq
tacraleln

CPI_0035 51146 49569
CT339 hypothetical protein
ARTILEEDAGSSLKPLEKTFPCATALYITHRRERKSEHOMMNRCQVFSSFFFRYPISSWL
IRLRASCECFQORHPIFLCGLYWLAGITSRGHPECSALILIFLGMELPRHPKOWLPLASA
WIISLMLTPAPFLLHGPISGTFVIHHAGGGGTYYGEALCIOTPCGKRAHHLSCGILSES
LELKKVZELECTLHHT3GIVFKSNACYKEIPRSFFVIMKEKCRESSCHFLNHRFPSSEVG
PFASSLLLCTPLPONLRDLFROKCLSHLFAISGWHFSLCATTLWMLCALLPLKIKKILSF
IVLTSLACIFPMSLSVWRSWISVTLLCFSWCFSGSCSGLNRLGAGFILCSIFFSPFSPTF
VLSFLATUGILLFFPKIFSFLYTFWTOFLSPFWLYPIR/LAMTLAISLSAQLFIVLPIMQ
YFGSLPLEGLLYNLIVPTILPIIVPLIATIILDCCSFITEALIGGFLSHPWLHNPNILK
TLSFAPVPPWMLTLA;LILFFIGILRTNVSPYASISATSYRFIETL

CPI_OUTS

ASSUMDSERKHMKIPDNOSTEDVRSFEPFOVLCTEQUPYEMSWEVVSAKTEPLPRGWYEL
ASSUMDSERKHMKIPDNOSTEDVRSFEPFOVLCTEQUPYEMSWEVVSAKTEPLPRGWYEL
ROSRCSFYSEPFULDPUDHRUPPLODRHYEKFFSTHIKFFSKWEDEGTFFMESLAKVVQK
LIKQULVVMNKMAEINKSSIASTEPFYGYEEPFAYQSFFFDDETRRULLSSHVLLNESGLE
HISCHEFTELLHLDSSYYPSFEJAHENYTHISEEPVNE

CHI_0037 51702 52115
pt3H PCC Phoophociti for Protein Hpr
kliskytelikat Lleisheinentetylegennertylegekdtodoffeloatitykn
aachyppasytyristen dynetyroktinaksimsilmlyapykdellyttrokea

HRILLOK LODAFSSGFGEL

Pen 0038 CTI 18 hypothetical protein
MDTOSSTCHERRIAGTSIVSCHALGKVFFLGTSPLHVRELTLPQEEVEHEIHRYYKALN
RGKSDIVALEQEVTOQOGLQEVGSILQAHLEIMKDPLLTEEVVNTIRKDRKNAEYVFSSV MGKIEEGUTAVRGMPGVVDRVQDIHDISNRVIGHLCCCHKSSLGESDONLIIFSEELTPS FVASANGAYIRGFVGLVGAATGHTAIVSRAKSIPYLANISEELWNIAKRYNGKLVLIDGY

EVASANIAT IN VOLUMENTATA VANDES PROBLEM SEEDIN AND RETURN OF THE PROBLEM OF THE P

NSCLEITEALLQAKTCSEVEELLNRNNKITS

CPn_0039 54256 53963
CT339 hypothetical protein
ISMOSGYAKKKEAKTHEQOFLEMEASLLEKRYEGQAGNGLVSVVINGKCDLISVKVQPT
CLDPEDPEVIEDLFRAAFKLAKEQMDQEMSLMRSTMPF

55673 54318

CPN_0040 55673 54318
dnaX-DNA POl III Gamma and Tau
AFYTHSLGYTMTLQPYQASSRKYRPQIFREILGQSSVVAVLKNALVFNRAAHAYLFSGIR
GTGKTTLARILAKALNCVHLSEDGEPCNQCFSCKEIASGSSLDVLEIDGASHRGIEDIRQ
INETVLTPPVVAKFKIYIIDEVHMITKEAFNALLKTLEEPPQHVKFFFATTEIHKIRGTI
LSRCQKMHLQRIPEKTILEKLSLMAQDDHIEASQEALAPIARAAQGSLRDAESLYDVVIS
LFPKSLSPDTVAQALGFASQDSLRTLDNAILQRDYATALGIVTDFLNSGVAPVTFLHDLT
LFYRNLLLTNSTTSKFSSQYKTEQLLEIIDFLGESAKHLQNTIFEQTFLETVIIHIIRIY
QRPVLSELISSIKSRQFFGLRNIKEPTLTQQVSAPQPQPTYKEQSFLEKKNQPAAEGKII
SVENKSSASIKSAAUTPILOFAVEFSGIHRO SVEVKSSASIKSAAVDTLLQFAVVEFSGILRQ

CPn_0041 55888 57342

No robust homolog present in Genebank/EMBL as of 11/7/98

CKYLYHHSYPPPQHSVGSISSRYKLRVLAITFLVLGVLLLISGALFLTLGIPGLTAGVSF

GLGIGLSALGGVLVVSGLLCLLVKREVSKVCPEEIPAVQPEETPBSVPVTPFEKPALDEA

QKEQKTOKILDQLPQELDQLDRYIQEAFACLGPLKDLKYEDQGFLQDVKEEFQVFDFVQK

DMIAEFVELQQILCQEGRLLEFVINGTRYIGRDLFKREDSLYKLWEWLGYLPSGDVRGER

LKKSAREVVDRFMRTTCNIRKIAMTFDRHVYSVAKTAFEKAFGALDTCVYESMRESYREA

FCEYEKAKLLGDEEKSAHAEORFQDIKNRWEDVKDAFFWVKEDGKIEIDDAIGNSCKWSE

RYEEHRITRARWYKVAEHQLFNATMRVKDSLREHNEARVAFEKERSKENQRQVQKKKEKR

LRÖBKELHDQELPRAQERLRELQALYPEIAVSVVEARREVASDLEKAHESIDKHYQSCVR

EÖEELY

LRDGKELHDQELPKAQERLEGASIFEANSTANDER
EQUITY

LITERATOR OF THE STATE OF DAFGAFQEMLKEGLVEEAQALRETEYWLYREERKSKKKH

CHIDO43 58432 60372

NO CODUST homolog present in Genebank/EMBL as of 11/7/98

HHRFTMQVPLSPQLPPPPPBHSVGASFCLSKFRVLAITFLVLGVLLLISGALFLTLGISG

VSLGVGLGLSALGSVLVISGFLLLLERREVSGVGLEGIPTGIPVGPSAEPSSEGIVAN

KNOILDQLPQELDQLDTDIQHVLSCLGKLKDLKCKDRGLLKDAKERLQVFDFVWKDMME

FVELQOVMDQESRYLEGIIHEVQSIAHKLFVDDVNIRSHLGESGGYLPSEDVRGELLKRF

AKEWARFMKVTRDIRKIAMAFNKNAYGAAKNAFDKAFGSLETCLYKSLTKSYRDTFCDY

KRAKILPDENNSARAEQAFREVKDHWEDLKETVFWVKEDGRIDIEVLTAVGGWPDRYPEL

LILETKRKDKVMSHOLWEATMRVKEAEVTYSVARVAFEKDGSQONQKKFQEKTKERLRCLK

DLRDGECHRAQERLEKLTALYPEVSVSVVETEREBKFNLEKAYGKLERYQSVVQDQEDY

WTEÖRNREAEFRAKGTKVRSMEEVAEHLQILENLLEDCYKKISKAETFALGVEREATEEI

EYTELSDAANRLKVLCEDIEDTLPRVEEIEMLRMAERPLHPIKQAFTKAFVQYNRCKER

LAKVEPYYKESPAYVNSEERLQSLDQASQCIQRVPKGFKFRNGSMYI

CPn 0044 60278 60778

CFI_2V44

Nokičobust homolog present in Genebank/EMBL as of 11/7/98

IANSDCRVWIRLHSAYKESGKVSSLETEACTYREYLREDVQOFETOGVSLIKEELLFLSS
TLKSKLSYDPLIANIPCMKFYYQYYDDIDKARAQSRWLEKSERYRNAKRRFQEIVKKGLF
KEARPLKKEEYRLLQEERSNKEKRLIYNKMAVARQRVQEFESMEIPE

CPh_0045 60961 62790
CT345 hypothetical protein
CKYTYHPPOLPPDHSVGATSWOPKLRILTITFLVLGVLLLISGALFLTLGVPGLAACLSF
GLGIGLSALGGVLVVSGLLFFLIRRGVSKVRPEEIPVTPSHEAQKILCQLPQELDQLDTS
IQEVVSCLGKLKDLKYEDQGLLTEVQEKLRVFDFVRKDMYTEFLELQQVVAQECQYLDYL
INQVGSISHKLFVPDVNIGAHLAELCGYLPSGDVRVERLKRSARQVVDRFMRVTDTRKV
AMAFDENACGVAKNAFDKAFGALBECVYKSLTESYPEAFFEYFEKAKILRHEDVSKLODKN
KSARAEQRFREVKDRWEDLKETVFWVKENGCIDLEVLTAVGGWPDRGPEHLIPEKRRNKV
MSHKLWEATMRMKGAEGTYSVARVAFEKKDSRKNQKKFOEKTKEWLRCLKDLJDQECHRA
RERLAELEALYPEVSVSVVETEREFKFKLETAYGNLEEFRYGSVVRÖEDFWYEEENKEAE
FREKGTKVRSPEEVVEYLQILENLSEDCSKQLTIAET/VLGVELEATAEFFYTILSDAAN
RLKVLCEDIEDILPRVEEIEIMLRIAELPFLPIKQAFTKAFLQYNSCKDKKLAKVEPYCQE
SVDYKSGFRV SVDYKSGFRV

CPn_0046 62775 63263
No robust homolog present in Genebank/EMBL as of 11/7/98
ERFOSLMODLONVYQECOKATGLESEVSAYROHLREGITEFETOSLIVIKEELLEVSSTL
KSKLS/OPLIADIPCMKFYEEYYDOIDKARVOSRMLEKSERYRKAKKGFQEMLKEGLFKE
DQALKKAEYRLLREKRMNKEKLLICNKIEAAQORVZEFGPSDS

CFM_0047

No fobilist homolog present in Genebank/EMBL is of 11/7/98
EMERILEVITYOFFREVLEGILITMYHFOKIRMTUTTG/F7LHKFLRKDYELMFVYGGCDES

CEN_0048 0.3687 65801
*MITE-BE CONSCIENT DYPOTHETICAL IN DYSTAIN
MKELPHEDYNKALIKLEHOWVRYFLYTFVOCHFIVALFTFAWLKVLYVPEYKAGEISRIS
IFARMDEGLEWEJNIKFYKKTAHIGEAFEKVYHLTUGFYJLLEKENADENTDYWFKKAAD
FILETTNFVDETFOKCIKDLCIYPPLLGKEKKTLEIHIJENKGNVIAQUFCHLKIFLIGEN

CPOPCEDA IMDILKIANIES ASSCVKGELLGKRCIEK ITKGTPILEKYORI DDRAKILKQLRAQLLSVHTLESS AIFVVLLILLWGYGALKALC PEMLKSPORFMLYIA ILTISLLWCRGTEI PCAYMORTES (PPILPFTAVLLGYFLG) PIAGFSCTFLALLYTLGS DLWNISWFLIS INLLCSWRILVSLHRVSRLSSVFWAGMKLOOVAMGSLLMFRIFTITISRE ALYADG IESFVYSLITAISVVALIEVFEASFGASTNFSLLTYLSPENALLKRLFKEAPGT YOHSVLVGSLABAAQA IGADSLYCLVAAHYHDJCKLINNGFFSENGKILQGSGHSLSPLECAKMIMRH IPEGVILARQAGLPESFIGVIELHGTSVIRAYYSHMVENFTGSFDEEL ECAKMIMRH IPEGVILARQAGLPESFIGVIELHGTSVIRAYYSHMVENFTGSFDEEL FRYSGNKPSSKETTI IMIADSFEAASRSLKNASLPDLQRLIDQI IQGKLQDGOFSCSPIT

No robust homolog present in Genebank/EMBL as of 11/7/98
LKEKRRNIVYLLVIYQEIFWLTMLHQPYYDKILTCMTIYIPGHTHKDSNKLFQKKSRAIW
VDEKPFSLDCFSNVFLIFVSLVPIAGLYRAYQIKKSLDRTTVQIGYSPSLSCEQKECVEA FVNGYGLICISILGGLGILVPILILVVLSLLLLGILMLFSLSTYESIKNYISKHICWKSN

CPn_0050 66849 66499
No robust homolog present in Genebank/EMBL as of 11/7/98
VSWFPILGIFLAMRYAKHOTNWNDENVKANLGYLPSTNCKNALCRMSSRLTSSIKTAGIL
GGGGILLPIFLLLLAILLISVLBQLIMLPFRLCCFALRGSVSSDTVTNLLLLINNTLA

66797 67111 CPn_0051

No robust homolog present in Genebank/EMBL as of 11/7/98 CFAYLIARNIPRMGNHETYIHPGVLPSSHAQDVSRSTVYPSRSFIMRRMLMGWNFNRVPS KSSEQLMDGHRIPLIFFGKHPTISILNVNRFSWLSIFYNGERGF

68008

CPn_0052 /68008 67304 hemC-Porphobilingen Deaminase whLsvCysDpcLsDfcQckRPLRIASRNSNLAKAQVHECISLLRSWYPKLWFQLSTTETT GDREKKIPLHLVENSYFFTDGVDALVHKGVCDLAIHSAKDLPETPSLPVVAITRCLHPAD LLVYADHYVHEPLPLSPRLGSSSLRRSAVLKQLFPQGQILDIRGTIEERLDQLHRGHYDA IVLAKAASLRLHLHAAYSIELPPPYHALQGSLAITAKDHAGKWKQLFTPIHCHSS

CPn 0053 69350

sms-Sms Protein
IRMATKTKTOWICHOCGATAPKWLGOCPGCHNWNSLVEEYVPQARSGTSSRSSTSAIALS
SIELENESRIFIDHAGWDRILGGGVVRGSLTLLOGDPGIGKSTLLLOTAERLASOKYKVL
YVCGEESVTOTSLRAKRINISSPLIYLFPETNLDNIKOQIATLEPDILIIDSIQIIFNPT
LNSAPGSVAQVREVTYELMQIAKSAQITTFIIGHVTKSGEIACPRVLEHLVDTVLYFEGN
SHANYRMIKSVKNRFGPTNELLILSMHADGLKEVSNPSGLFLOEKTGPTTGSNIPIIEG
SGALLIELDALVSSSPFANPVRKTAGFDPNRFSLLLAVLEKRAQVKLFTMDVFLSTTGGL
KIIEPAAOLGALLAVASSLYNRLLPNNSIVIGEVGLGGEIRHVAHLERRIKEGKLMGFEG
AILPEGOISSLPKEIRENFRLOGVKTIKDAIRLLL

0654 70089 69313

CPILJUS4 03313
ric_Ribonuclease III
rLSFPPPIKIPNSKFKDGALLSMHPPIDITAIEAKLNFTFTQPKLLEIALTHPSYKNESA
QQEDSERLEFLGDAVLGLIVTEHLFLLFPSMDEGTLSTARASLVNAKACCRYTTMLGIG
FWLLIGKGEKIOSERGRUSAVANLFESILGAVYLDGGLSPARKLTVPLLPPREEILPLMS
GMPPMLLQOFTOKOFRVLPVYQSTAVTDAQGNVSYQIQVLVNQEVWGEGNASSKKEAEKI
MAQQALDTYGNKNQNTMDV

0055 70096

70330 77296 hypothetical protein CFWICYLIRIRMRSALHLOHLRHFHNHGSILFENLLTIKDCFLLETKLONFIAKASKTID TVRWRENIFRSMPEIYTVVRKRRLDFFAAELVHRPKLSLVRDLWVFPGEEILEGEEDCML LLSGDRAGSGIFFTGPYPSDLYELEKGTTGLLLAFSSVGIPVI

70917 72746

CPn_0056 70917 72746

mrsa-Phosphomannomusse

EFLKLSLHRISLMKEVEQRIRSLYDAVTAENICRWLSNDCTQQDAKTILGWLDTDPAQLE

DLFGATLTFGTGGLRSLMGIGTNRINLFTIRRTTQGLVQVLRAHLPHPGDPMRVVVGCDT

RHNSIEFAQETAKVLAGNGCEVFLFQYEPELALVSFTVRYERRIGGVMITASHNPPNYNG

KVVMASGQVLPPLDQEIVAACSAVNEILSVPSIDHPNIHLIGKEYEALYRDTLKQLQL

YPEANRISGRSLSISYSPLHGTGISLVPHVLKDMGFLSVHLVEKQAIGDGDFTVVQLPNP

EDPEALTLGIEQMLANDDDLFIATDPDADRVGVVCLEDGQPYRFNGNOMASLLADHILGA

WSKTRHLGEHDKLVKSLVTTEMLSAIAKHYHVDLINVGTGFKYIGEKIESWRNSTNKFVF

GAEESYGCLYGTHVEDKDAIIASALIAEAALQKLQGKTLCDALLSLYETYGYFANKTES

VVFSAKTDEQEIRKKLSHLEEISSANFFSGKYQVEKFENYKQGIGFNLLSKDSYALTLPK

TSMLCYYFSGGGRVIIRPSGTEPKIKFYFEMSTHYPERVTDKEIQKQREAESFQHLDDFI

FDFKEKFSNL FDFKEKFSNL

CPn_0057 72913 73554

CPI_UO37 ... 3554
sodM-Superoxide Dismutase (Mn)
ILKRYVVMSFVPYSLPELPYDYDALEPVISSEIMILHHOKHHQIYINNLNAALKRLDAAE
TQQNLNELIALEPALRFNGCGHINHSLFWETLAPIDGCGQPPKHELLSLIERFWGTMDN
FLKKLIEVAAGVQGSGWAWLGFCPAKOELVLQATANQDPLEPLTGKLPLLGVDVWEHAYY
LQYKNVPMDYLKAFFQIINWGHIENRFSEIISSK

73627

CFI_0038

ACCD-ACCOA Carboxylase/Transferase Beta

IRMLVRLFSYNKPKIKVQKIKADGFSGWLKCNHCHEMIHANELGONYNCCPKCSYHYRIT

AIERVKLLADKDSWRPLYTDLKSODPLEFIDTDTYANKLEKARKNTTESEGVIVGICTIG

LHPVALAWMDFNFMASSMGAVVGEKLTPLIEEAIETRLPVIIVSASGGARMQESVFSLMQ

MVKTSAALAKLHEAGLPYISVLTNPTSGGVTASFAALGDIIIAEPKALICFAOPRVVAQV

IGEDLPEGAOKSEFLLEHGMIDKIVERKELKTTLQTLLDYFLAQEYTOGKSKAPRDLSKR

LKEIFLLTDDSF

74562

CPG_0659 74562 75050
dut-dUTP Nucleot idobydrolas;
tKHHTAJCNONITICNATEMTYFCELDCCCELPEYTTPGAAGADERANIEEPIALLPYQRA
LIPPNIKAEIPEGYELQVRPRSGLALKHGITVLHISPGTIDSDYRGEIRVILINFGDSTFI
IEPKMFIAQVVLGPVVQATFVVKQEGLAETARGCCGFGHTGAS

CPn_0650 75004 7552H
pcsh-PTS_IIA_Prote-in
PKLPEEYEVUTLEQAKMPSYCOMOODFOLFSLLSPRLVMFLGKHSRDETLQULTDLVDA
AGLLEDKQAFFDALVRREN INSTSTGWZVATPHOKLESCONFFTATGTHTQGTLWDATDG
ALVRLVFLTGGPENAQAEYLKLLSTLTLSLREESPROOLLQVNTTEEVMNVFVGM

CPN_0061 75501 76208
ptsN-PTS IIA Protein + HTH DNA-Bindin
RSHEC ICKDVMNDLKLDEVASLLDVSEHTVLOWLKEG. SMNNEYRFSREEIENWLL
HNQALM IQERCEDKEALKDLSLKYSLYKAIHRGGVLCDVVVHSKEEALQVASKYIAQKFQ
LDESVLFBMLHRENLMSTG IGEGIALPHAKDFLINAYYDIVVPMFLAEPIEYGALDGKP
VGILFFLFACODKSHLNLVNKIVHLGMSLNARSFFKNYPNKDQLLAYVKEWESQTH

CPD_0062 76251 77690

CTDS-chypotherical participate (avel) 1 PFTE CPANAGE (AVEL)

EVVKKPOKGSERKOAKKEPRARKGYLVPSSRTLSARAJKMKNSSRKESSCXCNEISANST
PRSVKLRNKRAEOKAARQGFSAFSNLTLKSLLPKLPSKOKTSIHEREKATSRFVNESQL
SSARKRYCTPSSAPSLFLETE IVRAPVERTKELODNEIH I PVVQVOTNPKEONTKTTKO
LASOAS IQOSEGTEOSLEELAGASLPVLVRSNPEV SVOROKEELLKELVAERROCKRKS
VRQALEARSLTKKVARGGSVTSTLRYDPEKAAE I KSRRNCKVSPEAREOKYSSCKEDARA
NGKODKTTPSEDASQEEQOTGAGLVRKTPKSQVASNAQNFYRNSKNTNIDSYLTANQYSC
SSEETDWPCSSCVSKRRTHNS I SVCTMVVTVIAMI VGALIIANATESQTTSDPTPPTPTP

CPn_0063 78109 78267 No robust homolog present in Genebank/EMBL as of 11/7/98 PMYANCKHNCLCLYDFSRHRSPPGLPLTFTPPYSFTLGIFLGRCLSTSNIVLL

78340 78576 CPn 0064 No robust homolog present in Genebank/EMBL as of 11/7/98 LVMTKIQCSAQYYRSRPAERAQTPPQPFLARDRADFWERHPRFSACCRVLLLVAWVVLAL LFLFVMLLPLAAGSYLLAF

CPn 0065 78882 80651 CT288 hypothetical protein
YDYYKYNMFFKKNYMTDFPTHFKGPKLNPIKVNPNFFERNPKVARVLQITAVVLGIIALL sgivliigtplgapismilogcllasggalfvogtiatiloarnsykkavnokklseplm erpelkaldysldlkevwdlhhsvvkhlkkldlnlsktorevlnoikiddbgpslgecaa MISENYDACLKWLAYREELLKEQTQYQETRFNQNITHRIKVLISIISHITDHISKAGGYA MISENYDACLKWLAYREELLKEQTQYQETRFNQNITHRIKVLISIISHITDHISKAGGYA SLKFSTLSSRMSRIHTTTTVILALSAVVSVMVVAALIPGGILALPILLAVAISAGVIVTG SLYLVRQILSMYKRNGODFYKDFVKNVOIELLAQVITUQFILFBLKGVLKEEEVSLEG ODWYTQYITNAPIEKRLIEEIRVTYKEIDAQTKKMKTDLEFLENEVRSGR.SVASPSEDP SETPIFTQCKEFAKLRRQTSQNISTIYQFDNENIDPEFSLPWMPKKEEEIDHSLEPVTKL EPGSREELLLVEGVNPTLRELNMRIALLQQLSSVRKWRHPRGEHYGNVIYSDTELDRIQ MLEGAFYNHLREAQEEITQSLGDLVDIQNRILGIIVEGDSDSRTEEEPQE

CPN_0066 80916 82655

No=robust homolog present in Genebank/EMBL as of 11/7/98
CYMANPTOSEPPSPEISIEELELOELAGSSNTETISNTPPPSCAATAEEVSLFIEGGRR
NSEDEEGPLGSCEV/DVVCITNGGDPEVRDHEVRWYINGSGRTOHEGILDAMNICDLRG NSEDEEGFLGSEVYDVVCITNGGDFEVRDHEVRWYYINGSGRTOHEGILDAMMICDLRG
EBYRFIRNSGYGLGSCFLGIRNRIPPRDNVISQA IQARWHEFIFAENANDDY VLFSGN
GGLYLOVALDNSIYSHHILCVGIGSSYYIQGNYRVHNYRVTGDWTTLLDRRGATAVNTTT
LIMADSAEGLFLPSVRCPSYWALKCGGCLIMDNNQVGFRPODGSSEIALVWNLNODH
ESTATRLIEW IDRGDSQAVLELNPOPSHCRDIALTALYATTRISSLLQBCLMISTYTAPEY
FARVAIVTGYSIMTLRYFILLLTNRRGCRHFRVLRLAALGLQSLGFLTVLLDHTNYTRR
WIRRPPLISVIFCTASFATGSFIYVDLTRWFFTSLRSRLQLFVQRRLTGRGLFLRRYFN
HLDSLRFSQNALITFHGGLFMPLIIGFFNQLVIQVPRVVIRPMTTAVYDLNOTSQEAMDS
GDWLAIGQTINFLLCMILLVINTFFFVRSVRRNLHRRPHR

CPTLO067 82920 84053
NO TODUST homolog present in Genebank/EMBL as of 11/7/98
KGSGYSYRGPPMAVEGRVNISSOALNODCOEVLANKOSKGLLRCRILSIVVAVITÄTIAGVV
LIALTILASILTSVPYLALGVFLLIVTLGCIIFALCSEKIKKVPPTPISHKEEIIWFEBR
KNEDMEKEKEDPEHFGRTATDIPMRSALDOFNHSCHHIHESPALTETYRSHODVLLPKDW
CPVTLPDVTSEEEVLIRSVVGSYLLMEACVPKVSMLIDELHNYLKEPSERECLFTKKTL
QRKASFLETOKDLATTFLAYTRVNDCHLAPFRAGAKWILIHYVALRRQHNONDFTFTGCO
CYYARLAFNOTORLYHQLFNVEKLRSIYANMDKDPLCHPWAFIPIYDLLKTEDHGDGFLE
OGERBEFYBRAAONDFWG QOEDREY PSRAAQDQFWG

84909 84331 CVT_U068 84909 84331
CTJ#50 hypothetical protein
SF#IKKFIYSLIFSCSFSAPLKGICNEDVSSQSRIEEDPEVLITQLNELIETPIEEGKE
IRRELOAISDCQKSSEEIEESCGTSDSEGLSEKTDKESSNEYVLDFFDSMVQRLEGISKM
CQSEQVAQIIDCFNREFDIRNRELELKNRELELREKDLEFKKSILDWYKEKVSRELAFQR EODIKOTLMLLKK

CPI_0069 85191 87086

No robust homolog present in Genebank/EMBL as of 11/7/98

No robust homolog present in Genebank/EMBL as of 11/7/98

NEFLYVYLLIFNLGIMTTPPPSRSSSPPPYDWIELQDLGNTNNNSSRATPPPPEVGGELP

PYFSASNFVVIERGAPSLPSPQOLLSLPEVSROPPPGYFDETASITSRTSEEMFGTLVST

LCCPANSERDWEDHEVNCIYIASTSDTQLEAVQGGMHITELRGEPVRVLVETCHLYAFAR

ENTCHSRLEVSHTVRAMTYFWDRFFSRMWNVGRRFLVFYGGNGGAYVQAALDSSMHTODI

YVLGLSPTVYIRGNYHVOHYRVRGFWPSCLDSLAACAENTSVLPYGGSSDGIFYPSLFSH

TFDNAIRYGERCLLVCSEDMGMLPETQQQTSPLTSLEGGHEVALVLNPQONPEALSIASR

LMHEERCGRLESNYMPGRSSNFFMTSMYVLVRLNTLAQIYLMSPYYSFGSNDIVCLIFIS

SAAVETVGYIFLTVTDSTCGRRYLRVPRLVCTGLRNLALBTTLLELLILSYPRSVEGVPF

NVRFILGWTCTTRVYFFANNLILHWFFRCLRRIGQLFVH/SIIGHTLGARITDLTLASMR

YAIVFPSIVSSCLLTALAHANTNILALDPYRLIESGDLJRPAFNDDEMQQADNPWDAYSI

GLVINTCIYMLILFANLIFMYSVRRYHRSRR 85191 87086

CPn_0070 87399 87208
No robust homolog present in Genebank/EMBL as of 11/7/98
YKVGLFHLKNONFFSNOSRTYEORFPKVSPHFESIJRLOSVGFSSOGTLLISFRDTELKR

CPn_0071 88066 87599

YT125 hypothetical protein
_KKOLROLLERICELOHARCLKKOHKIIEELERERFOKDHLYLKLMENSSSRDAFDKKRMLT
KENLVOKOODLYLYEVYODGILFFFTYTKALMSSGIASLFTEVYSGETPSTILTCKPIF
FORLTPYLHFORLMKHESLYMRMKOIAVOYLKPPOT 88066 87599

89151 88067 CITE 2007: 89151 MND07
CITE 2007: B9151 MND07
CITE 24 hypothetical protein
RGYRSTKTEVTKEKVLLLIYGLLEFYFFHYMMSTPLSSGGISPSDQYVPQELFCDRLSSSR
SNSFDSHARSGGSPIVSPPISALVALTDLKLVPYNGNSFSWTTRLKNAVEKIGLFLQRNWK
YILLYILAWALTLYCHRTVALTLTIWLAVGLGIGVVFGIFTATCLDKENHRHVNSLWNL
INNGILUDDPWTTRQILLATMIASISALIYAVPQAVGLVIGFSIGNQLSINTVYGARLGD

iq i i kiigm i nokqunal i e i nrnngtdpatanllas LDLNNNSPDD i i aradqc imtlsgtlqq i kkepdr i EATYA LORKAHKKP LEN L LKLINLINGPMPYCFCMPEC

CPn_0073 / 89353 infA-Initiation Factor IF-1 89574

SMAKKEDTLVLEGK/EELLPCMHFRVTLENCMPVTAHLCGKMRMSNTRLLVGDRVTVEMS AYDLTKARVVYRAR

90955

LUTA-ELONGATION FACTOR TU
EUTA-ELONGATION FACTOR TU
EDFEMSKETYORNKPHINIGTIGHVDHGKTTLTAAITRALSGDGLASFRDYSSIDNTPEE
KARGITINASHVEVETPNRHYAHVDC PGHADVVKNMITGAAQMDGAILVVSATDGAMPQT
KEHILLARGVGVPYIVVFLNKVDMISQGDAELIDLVBEMELSELLEEKGYKGC PIIRGSAL
KALEGDANYIEKVRELMQAVDDNI PTPEREIDKPFLMPIEDVFSISGRGTVVTGRIERGI
VKVSDKVQLVGLGETKETIVTGVEMFRKELPEGRAGENVGLLLRGIGKNDVERGMVVCQP
NSVKPHYKFKSAVVVLQKEEGGRHKPFFSGVRPQFFFRTTDVTGVVTLPEGTEMVMPGDN
VERLMER IGTUALEEGGREAFBEGGRHKPGISKYNA VELDVELIGTVALEEGMRFAIREGGRTIGAGTISKINA

91087 91350 SRESFERKQOHNRKALSRKIGTVKKQAKFAGSFLDEIKKIEWVSKHDLKKYIKVVLISIFG FGFAIYFVDLVLRKSITCLDGITTFLFG

n 0076 91334 91903 Cyn_0076
yisG-Transcriptional Antitermination

PPCSVNCMYKMYVVQVFTAOEKKVKKALEDFKESSGMTDFIQEIILPIENVMEVKKGEH

KVVEKKIWROYLLVKMHLTDESWLYVKSTAGIVEFLGGGVPVALSEDEVRSILTDIEEKK

SGVVQKHQFEVGSRVKINDGVFVNFIGMVSEVFHDKGRLSVMVSIFGRETRVDDLEFWQV **EEVAPGOESE**

CPn 0077 91956 92435 T111-L11 Ribosomal Protein
FFVSYPLFVEVSQCKVRFSMSVKKVIKIIKLQIPGGKANPAPPIGPALGAAGVNIMGFCK
EFNAATQDKPGDLLPVVITVYADKTFTFITKQPPVSSLIKKTLNLESGSKIPNRNKVGKL
TQAQVEAIAEQMHKMDIVLLESAKRMVEGTARSMGIDVE

CPn_0078 92453 93160 T11-L1 Ribosomal Protein
SCRIMTKHGKRIRGILKNYDFSKSYSLREAIDILKQCPPVRFDQTVDVSIKLGIDPKKSD
QQIRGAVFLPNCTGKTLRILVFASGNKVKEAVEAGADFMGSDDLVEKIKSGWLEFDVAVA
TPDMREVGKLGKVLGPRNLMPTEKTGTVTTDVAKAISELRKGKIEFKADRAGYCNYGVG KLSFESSQIKENIEALSSALIKAKPPAAKGQYLVSFTISSTMGPGISIDTRELMAS

93170 93688 Tilo-Lio Ribosomal Protein
RGKNKQEKTLLLOEVEDKISAAQGFILLRYLRFTAAYSREFRNSLSGVSAEFEVLKKRIF
FKATEAAGLEVDCSDTOBLLGVVFSCGDFVSAAKQVLDFNKQHKDSLVFLAGRNDNASLS
GAEVEAVAKLPSLKELRQQVVGLFAAFMSQVVGINNSVLSGVISCVDQKAGKN

CPn_0080 93720 94 rl7-L7/L12 Ribosomal Protein VRVTKVTTESLETLVEKLSNLTVLELSQLKKLLEEKWDVTASAPVVAVAACGGGEAPVAA EPTEFAVTLEDVPADKKIGVLKVVREVTGLALKEAKENTEGLPKTVKEKTSKSDAEDTVK KLODAGAKASFKGL

CPn_0081 94219 98016

rpob-rna polymerase beca

freilshonsrrtrhlkcpervsvkkkedipdlpnlieiqiksykofloigklaeereni
gleevfreifpiksyneatvleylsynlovpkyspeecirrgitysytlkvrfrldetg
gleevfreifpiksyneatvleylsynlovpkyspeecirrgitysytlkvrfrldetg
gleevfreifpiksyneatvleylsynlovpkyspeecirrgitysytlkvrfrldetg
gleevfreifpikofffiingaervvvsovhrspeinyfecekskolifesfriipy
rdswileaifdindliyihidrkkrrkilaitfiralgyssdadiieefftigesslrse
kdfallvgriladniideasslvygkageklstahlkrrldagiasykiavdadehhpii
kmlakdpytosyeaalkdfyrrlregepattlanarstimrlfpppkrynlgrogvryklnrk
lofsiddealsovtlrkedvigalkylrrlkmodekacvddidhlanrvrsvgeliono
crsglarrkivpermilfdfssdtitpckvvsakglasvlkdffgrsolsofmodtnev
aelthkrrlsalgpgglareragfevrdvhashygricpiefpegpniglitslssfaki
nefgfieffyrrivrgivtdidifylethddanralmgsnmorqavpllkteapvvg
reflethyrtvthmdvspkolvsivtolifplehddanralmgsnmorqavpllkteapvvg
rglecraakdsgaivvabedsvvdfvdgykvvaakhpptikrtyhlkkflrsnsgttin
ogpleavgdvitkgdviadgatdrgelalgknvlvaprpwygymfedaiiiseklired
aytstyleffeltaaptrikigeeitridipnvsdevlandgeoiirigaevkpolilvek
itpksetelapeerllraifgekaadvkdasltvppgtegovmdvkvfsrkdrlsksdde
lveeavhlkdlokgyknovatlkteyrekelgalllnekapaaiihrrtaetvhegllfd
octierieogelvollmencemyevlkgllsdyftalgreinyktevelregodadldh
gvirovkvyvaskrkloyddmagrhgnkgvsxivpbadmylsngetvomilnplgvp
srmligovlethlgyaaktagiyvktpvfegfpegriwdmieoglepedgrsflydgktg
erfdnkvvigyiymlklshliadkiharsiopyslvtogorggamgorfgemewal
eaygyahmloeiltvksddvsgrtiyesivkgenllrsgtpesfnvlikemoglgldvr
pmvvda

CPI_0082 97992 102221

rpoC-RNA Polymerase Beta'
CSSYGRRRLKNDVLEKIMFGENSRDIGVLSKEGLFDKLEIGIASDITIRDKWSCGEIKKP
CSSYGRRRLKNDVLEKIMFGENSRDIGVLSKEGLFDKLEIGIASDITIRDKWSCGEIKKP
ETINNRTFKPEKGGLFCEKIFGPTKDWECCCCKYKKIKHKGIVCDRCGVEVTLSKVRRER
MAHIELAVPIVHIWFFKTPSRIGNVLGMTACDLERVITYEEYVVIDPGKTDLTKKQLIN
DAOYREVVEKMCKDAPVAKMCGEAIYDLLKSEDLQSLLKDLKREKKTKSQQARMKLAKR
LKIIEGFVSSSNHPEWMYLKNIPVYPPDLRELYPDLDCGRFATSDLNDLYRRVINRNNRLK
AILRLKTEEVIVRNEKRHLQEAVDALFDNGHCHPWMGAGNRPLKSLSEMLKCKNGRFRQ
NLLGRRVDYSGRSVITVOPELKFROCGIPKENALELFEPPI I KRIKLDQOSVYTIRSAKMI
IORGAPEWDVLEGIIKGHPVILNFAPTLHBLSIOAFEPVLIEGKAIRIHPLVCAAFNADFODDOMANIVPLSVEAQLEANVLMMAPDIHFLSIOAFEPVLIEGKAIRIHPLVCAAFNADFODDOMANIVPLSVEAQLEANVLMMAPDIHFLSIOAFEPVLIEGKAIRIHPLVCAAFNADFODDOMANIVPLSVEAQLEANVLMMAPDIHFLSIOAFEPVLIEGKAIRIHPLVCAAFNADFODDOMANIVPLSVEAQLEANVLMMAPDIHFLSIOAFEPVLIEGKAIRIHPLVCAAFNADFODDOMANIVPLSVEAQUEANVLMMAPDIHFLSIOAFEDCHGRIFIHEKIRIFOTOLIETT
GRRVLFIRITVEKELFOMYSMMCHEIGELIA/YKKVGLEATVRFLIDDLKDLGFIQATKA
ALSMCIKOVETSDIKSHILKDAYVLVAAIVKKVOLEATVVRFLIDDLKDLGFIQATKA
ALSMCIKOVETSDISSHGARKGLADTALKTADSGYITPRLVDVAQDVITPREKXSTLNHETEISAI
COCSEELLPLKORTYGRTVAEDVYOMYSMINDELBIRGGGDVINSVQARAIDDASIETTIKISI
COCSEELLPLKORTYGRTVAEDVYOMYSMINDIGENTGLITAADSGOTTHRETHILXITAATS
TUTGELFREGOTLVVMODLRVVILOAGENININIAIRIKGIGCHETAALHVVGDEGRTELNEYKKLLSTKGIE
SLEVPEVELGVKILVADOTTPVSXXXIIGEVELBHIPIICOKNOFIKKELLSTKGIE
SLEVPEVELGVKILVADOTTPVSXXXIIGEVELBHIPIICOKNOFIKKELLSTKGIE
SLEVPEVELGVKILVADOTTPVSXXXIIGEVELBHIPIICOKNOFIKKELLSTKGIE
SLEVPEVELGVKILVADOTTPVSXXXIIGEVELBHIPIICOKNOFIKKELDLWGGGTEKKV CPn_0082 97992 102221

NKNTGLVEL IVKQHRGELHPQ IA IYDDADLSELVGTY ISVEEGORVDPGMLLA RLPRGAIKTKDITGGLPRVAELVEARKPEDAADIAK GMEEEHLIPLTKHLIVQRGDSVIKGQQLTDGLVVPHE KGIQKNKRILVVCDEMT GVRELQKYLVNEVQEVYR LOGVDINDKHIEIIVROMLOKVRITDPGDTTLLFGEDVNKKEFYEENRRTEEDGGKPAOA VPVLLGITKA:LOTESFIGAASFODTTRVLTDAACCSKTDYLLGFKENVIMGHMIPGGTG FETHKRIKOYLEKEQEDLVFDFVSETECVC

102296 CPn_0083 103312

Torareald for a exemplative monopropertion every objective accessorative of lenvageer f QELLNEAVVWG IRQNGDDLOTLSFILDK IQVNFALE I IKNI FOR I SLE IDARLSFNVEAM VQRAVFLSQLFEAMCGDKKRLLVK I PGTWEGIRAVEFLEAKGIACNVTLI FNLVQA I AAA KAKATLI SPFVGRI YDWWI AAYGDEGYSIDAD PGVASVSNI YAYYKKFGI PTQ IMAASFR TKEQVLALAGCDLLTISPKLLDELKKSOHPVKKELDPAEAKKLDVQPIELTESFFRFLMN EDAMATEKLAEGIRIFAGDTQILETAITEFIKQIAABGA

103356 103751 predicted ferredoxin

SEMKNKMDYKSQLVFSCPCCKGNVCFSVFNLDVILTCNVCSSTYTFDSVIRNEIRQFVA LCKRIHDANSILGNATVSVSVEDNQMDIPFQLLFSRFPVVLNLSLDGKKIAIRFLFDALN TSILHOESDLIS

104512 103766 CPn 0085

CPI_DU053

CTJ11 hypothetical protein
FSMKFFILFILIVAOFPAFSAQPRTOVSASHSKQAKARRTSRIRSSAATNASVSRYKTRA
AARKKIGKFEKKPSLSPVQWYRYSGKNYSIQTPSLWQCIDDKTQLPEKLDVLLIGKGKGN
LTPTINIAQEITSKSSKEYIEEILAYHKANEWTLESGIFTQIQSPSGEFTIIKTEKNSSW GRVFCLQATTVIDHTAYIFTSTATLDDYAELSFTFLKVVSSFQIRGGKEATSGDAILEKA

CPn_0086 104898 atpE-ATP Synthase Subunit E 105527

acpe-alp Synchase Sudditt E NIMANLANDGKLKOJCDALRLDTLKPAEDEAAALLHNAKEQAKRIIQEAQEEARKILETA EERAHOKIKOGEVALSQACKRALEALKQAVENKIFRESLVEWLEHVITDPEVSTKLIQAL VQALEAQGVSGNLTAYIGKHVSPRAVNELLGKAVTTKLRKKSVVVGSFVGGVQLKVEEKN WVLDLSSSALLEIFTRYLQKDFREMIFQGS

105540

CPT_0087 105540 106376
CT309 hypothetical protein
SHEKIFSIFKVVMTQYYFLSSFLPTQLPESVPLFSISDLDDLLYLNLSENDLCNYGLLK
RFPDFENFAFFMAGKPIPFSFGEVTQENVERMLSSQQMSDDNDFEDFFKDFLMNHKSSQD
REMHFSDLFREFLSYHQTNSSKFLQDYFRFQQQLRVVLAGFRARVLNMDVSYVLRDEDSS
DEWLEVLMQKDSPNYELPEEFSDLGGVLDDYGLLPHTLNRALALYQFHKLEGFCSDSYF
DEWLEVLMQKDSPNYELPEEFSDLGGVLDDYGLLPHTLNRALALYQFHKLEGFCSDSYF
DEWLEVLMQKDSPNYELPEEFSDLGGVLDDYGLLPHTLNRALALYQFHKLEGFCSDSYF
DEWLEVLMQKDSPNYELPEEFSDLGGVLDDYGLLPHTLNRALALYQFHKLEGFCSDSYF

THE OOS B 106352 108145

CT-288 hypothetical protein
SYRKCHOMYTVSEOTAOGHVIEAYGNLLRVRFDGYVRGGEVAYVNVDHTMLKAEVIEVAD
DELWOVYEDTOGACRGLVTFSGHLLEAELGPGLLOGIFDGLORRLEVLAEDSSFLORG
KHTNAISDHNLWNYTPVASVGDTLRRGDLLGTVPEGRFTHKIMYPFSCFOEVTLTMYTSF
GTYNAHTVVAKARDAOGKECAFTMVORWPIKOAFIEGEKIPAHKIMDVGLRILDTOVEYL
KÖGTFCTPOFGGAGKTVLOHHLSKYAAVDIVILCACGERGEVVEVLOEPHLIDENSYL
KÖGTFCTPOPFGAGKTVLOHHLSKYAAVDIVILCACGERGEVVEVLOEPHLIDENSYL
SÜFELPGEFAFPAYLSSRIAAFYERGGATTKDGSEGSLTICGAVSPAGCHFEEPYTGST
LAVVGAFCGLSKARADARRYPSIDPLISWSKYLNOVOGILEEKVSGWGAVKAADFUST
GSEIGKRWEVVGEEGVSWEDMEIYLKAELYDFCYLONAFDPVDCYCPFERQIBLFSLIS
RIFDAKFVFDSPDDARSFFLELQSKIKTLNGLKFLSEEYHESKEVIVRLLEKYMVOMA

108111 109466

CT289 hypothetical protein LBCWKKOWYKWRKDMQTIYTKITDIKGNLITVEAEGARLGELATITRSDGRSSYASVLRF DEKKYTLOVFGGTSGLSTGDHYTFLGFNEVTFGSSLLGRRINGIGKBYDNEGECFGEPI
EIAPPTENPVCRIVPRSMYRTNIPMIDVFNCLVKSQKIPIFSSSGEHMALLMRIAAOTD
ADITVIGMGLTFVDYSFFVESKKLGFADKCVMFIHKAVDAPVECVLVPDMALACAEKF
AVEEKKNVLVLLTDMTAFADALKEISITMDQIPANRGYPGSLYSDLALRYEKAVEIADGG
SKILITYTTMPSDDITHPVPDNTGYITEGOFYLRNNRIDPFGSLGRLKQLVIGKYTREDH
GDLANALIRLYADSRKATERMAMGFKLSNWDKKLLAFSELFETRLMSLEVNIPLEEALDI
GWILLADSPSFSFVERAR LANDVARDER GWKILAQSFTSEEVGIKAQLINKYWPKACLSK

CPn_0090 109439 110080
atpD-ATP Synthase Subunit D
VLAKSMSVQVKLTKNSFRLEKQKLARLQTYLPTLKLKKALLQAEVQNAVKDAAECDKDYV
QAYERIYAFAELFSIPLCTDCVEKSFEIGSIDNDFENJAGVEVPIVREVTLFPASYSLLG
TPIWLDTMLSASKELVVKKVMAEVSKERLKILEEELBAVSIRVNLFEKKLIPETTKILKK IAVFLSDRSITDVGQVKMAKKKIELRKARGDECV

110074 112053

CPD_0091 110074 112053
atpI-ATP Synthase Subunit I
VRLNIHKYLFIGRNRADFFSASRELGVVEFISKKFITTEOGHRFVECLKVFDHLEAEYS
LEALEFVKDESVSVEDIVSEVLTLNKEIKGLLETVKALRKEIVRVKPLGAFSSSEIAELS
RKTGISLRFFYRTHKDNEDLEEDSPNVYFLSTAYNFDYYLVLGVVDLPRDRYYEIEAPRS
VNELQVDLANLQREIRNRSDRLCDLYAYSBEVLRGLGNYNDEORLHQAKECCEDLFDGK
FAVAGWVIVDRIKELGSLCHNYQIYMENØYVDPDETIFTYLENKGVGMGEDLVQIYDTP
AYSDKDPSTWVFFAFVLFFSMIVNDAGYGLLFLMSSLLFSWKFRRKMKFSKHLSRMLKMT
AILGLGCICWGTTTTSFFGMSFSKTSVFREYSMTHVLALKKAEYYLQMRPKAYKELTNEY
PSLKAIRDPKAFLLATEIGSAGIESYVVYDKFIDNILMELALFIGVVHLSLGMLRYLRY
RYSGIGMILFMVSAVLYVPIVLGTYSLHYLFHVPYELGOGIGLAVVLAMI
QRUMRGVEEIISVIQVFSDVLSYZRIYALGLAGAMMGATFNQMGARLPMLLGSIVILLGH
SVHIILSIMGGVIHGLRLNFIENYHYSFDCGGRPLRPLRKIVCSEDAEASGIHLDNNSIV

HOSSIYOKCYAA IGIVESTSLEAVVEALLLL

112440 113015

CT-003 hypothetical protein CT-03 hypothetical protein CKACVVIAEEKIM DOLGAN MOREGUSMOREGISHIFHCELEFLRYYSKEVFGUTVELAA GSV CHARVVIAEEKIM DOLGAN MOREGUSMOREGUSMOREGUSMOREGUST CHARVVIALCSONOET VSRH LYTHWYYYGACKVEKELEFYELGGI CHYRVALCSONOET VSRH CYTHWYYYGACKVEKELEFYELGGI CHYRVALCSONOET VSRH

HEWMEY ESLECT

11597/

CPn_0094 113104 11597

vals-valy1 erna synthetash
Vyrvflsrdhkfglrimttedfpkaynfodtepel/vfwekngmfkaeassdkppysvim
pppnytgvlhmchalvntlodvlvpykryggfevcwindtdiagiavavkafkt
lefengyiyrgyvlvmdpdvlotaladdeveyeekd mytdiagiavaggesivvattrpe
The analymentalvatlodvlotaladdeveyeekd mytgyrrygggsivvattrpe
The analymentalvatlogvlotaladdeveyeekd mytgyrrymgggsivvattrpe
The analymentalvatlogvlotaladeveyeekd mytgyrrymgggsivvattrpe
The analymentalvatlogvlotaladeveyeekd mytgyrrymgggsivvattrpe
The analymentalvatlogvlotaladeveyeeld valed helpingressory
Vivgyrsgaviepylskumpvovy/analymentalverdevilskipfkdpvkytldwnhlrdw
Cisrolmychripuskympvovy/analymentalverdylkipfkdpvkytldwnhlrdw
Cisrolmychripuskympvovy/analymentalverdylkipfpswydppovldtmfssglmp
Ltclgmpdenspblkkyyptalvytghdilffwytrmvllssmsgekpfsevflhglif
Gksyrryndfgewsiskkekdaydmgealpovakweklskskgnvidplemiatygt
Davrltlgscanregidldyrfeeykhfankwngarfifghisdlogkbllagided
Slgledfyildofnglihgyeeavatyafdkyatlayeffrndlgalpeddgdaftghalr
MLRSRacmeapypkafdykipgdlresftlagrlvytirnirgenoldprihlkafyvcs
Dttelgscipiloalgalesiglldkepekglysgvvdtirlgifypeehllkergrle
Kervrleravenlerilgdesfcokanpnlvvakgealknnrielggildklasfa

CPn_0095 115956 118790
pknD-s/T Protein Kinase
ACIVCLDREDOSLERYDIVRIIGKGGMGEVYLAYDPVCSRKVALKKIREDLAENPLLKR
RFLREARIAADLIHPGVVPVYIIYSEKDPVYYTMPYIEGYTLKTLLKSVWQKESLSKELA
EKTSVGAFLS/FHKICCTIEYVHSRGILHRDLKPDNILLGLFSEAVILDWGAAVACGEEE
DLLDIDVSKEEVLSSRMTIPGRIVOTPDYMAPERLLGHPASKSTDIYALGVVLYQMLTLS
FPYRRKKGKKIVLDGQRIPSPQEVAPYREIPPFLSAVVMRMLAVDPQERYSSVTELKEDI
ESHLKGSFKWTLTTALPPKKSSSWKLNEPILLSKYFPHLEVSPASWYSLAISNIESFSEM
RLEYTICKKGLNEGGGILLPTSENALGGDFYQGYGFWLHIKERTLSVSLVMSLEIQRG
QDLESCKETFLIALEQHNHSLSLFVDGTTWLHHMYLPSRSGRVAIIVRDMEDILEDIGI
FESSØSLRVSCLAVPDAFLAEKLYDRALVYRRIAESFFGRKEGYEARFRAGITVLEKAS
TDNYEGGFALAIEEFSKLHDGVAAPLEYLGKALVYQRIQEYWEEIKSLLLALKRYSQHPE
IFFLKDHVVYRLHESFYKRDRLALVFMILVLEIAPOATTPGGEEKILLALKRYSQHPE
IFFLKDHVVYRLHESFYKRDRLALVFMILVLEIAPOATTPGGEEKILLALKRYSQHPE
IFFLKDHVVYRLHESFYKRDRLALVFMILVLEIAPOATTPGGEEKILVWLKDKSRATLFC
LEDPTVLELRSSKMELFLSYWSGFIPHLNSLFHRAWDQSDVRALIEIFYVACDLHKWQFL
SSCIDIFKESLEDCKATEEIVFFSFEDLGAFLFAIQSIFNKEDAEKIFVSNDQLSFILLV
XIFDLFANRALLESGGEAIFQALDLIRSKVPENFYHDVLRNHEIRAHLWCRNEKALSTIF
ENYTEKGLUREOHEHFFLYGGYLALIQCAEAAKOHFDVCREDRIFPASLLARNYNRLGLP
KDALSYQERRLLLRQKFLYFHCLGNHDERDLCQTMYHLLTEEFQL 115956

124347 CPn_0096 118837

CPn_0096 124347 118837
CT296 hypothetical protein
ETFLSILREFFMKSLPVYVSGIKVRNLKNVSIHFNSEEIVLLTGVSGSGKSSIAFDTLYA
AGRKRYISTLPFFFATTITTLENPKVEEIHGLSPTIAIKQNHFSHYSHATVGSTTELFSH
LALLFTLEGQAADPKTKEVLDLYSKEKVLSTIMELSEGVGISILAPLLRKDIAAIHEYAQ
GSFTKVRCMGTHPLYSFLTSGIPECSVDIVIDTLIKSENNIARLKVSLFTALEFGEGH
CSVLSDEELMTFSTKQQIDDVTYTPLTQQLFSPHALESRCSLCQGSGIFISIONPLLIDE
NLSIKENCCSFAGNCSSYLYHTIYQALADALMFNLETPWKDLSPEIQNIFLRGKNNLVLP
VRLFPOTLGKKNLTVKWRGVLNDIGDKVRYTTKFSRYLSKGMSAHSCSLCKGTGLGDYA
SVATWEGKTFTEFQOMSLNNMHVFFSKVKSPSLSIQEILQGLKQRLSFLIDLGLGYLTFN
RALATLSGGEQERTAIAKHLGGELFGITYILDEPSIGLHPODTEKLIGVIKKLRDQGNTV
ILVEHEERMISLADRI DIGFGGAIFGGEVLFNGKPEDFLMNSSSLTAKYLRQELTIPIP
ESREAPTSWLLLTEATIHNLKNLSIRLPLARLIGVTGVSGSGKSSLINNTLVPAIESFLK
QENFKNLHFEMGCIGGIHTTRDLFGRSQRSIPLTYIKAFDDIRELFASQPRSLRQGLTK
AHFSFNOPGGACIGCGGLGTMTISDLDDTPIPCSECQGKRYHSEVLEILYEGKNIADILDM
TAYEAEKFFISHPKIHEKIHALCSLRLDYLPLGRPLSTLSGGEIGRLKLAHELLFASPKQ AHFSRNOPGGACIQCGLITMTISDDDTPIPCSECQGKRYHSEVLEILYBCKNIADILDM
TAYEAEKFISHPKHEKIHALCSLRLDYLPLGRPLSTLSGGEIQRLKLAHELLFASPKQ
TLYVLDEPTTGLHTHDIQALIEVLLSLTYLGHTVLVIEHMHVVKVCDYVLELGPEGGDL
GYLLASCTPKDLIQLNTPTAKALAPYIEGSLDIEVVKSEPPSSPKSCDILIKDAYQNNL
KHIDLALPRNSLIAIAGFGASGKHSLVFDILVASGNIAVAELFPPYIRQGLLKETPLPSV
GEVKGLSPVISVRKCSSSNSSYHTIASALGLSKOLEKLFAILGEPFSPLTEEKLSKTTPQ
TIIDSLLKSYKDDYVTITSPIPLGSDLEIFLQEKQKEGFIKLYSEGNLVDLDERLPLNLI
EPAIVIQHTKVSPKNSSSLLSAISVAFSLSSEINTYISQKKQRKLSYSLGWKOKKGRLYP
EITHQLLSSDHPEGRCLTCGGRGEILKISLEEHKEKIAHYTPLEFFSLFFPKSYMKPVQK
LLKDENASQPLKLITTKEFLHFCRGSSEFPGKNALLMEQLDTESDSPLIKPLLALTSCPA
CKGSGLNDYANYVRINNTSLLDIYQEDATFLESFLNTIGTDDTRSIJQDLMNRLTFISKV
CKGSGLNDYANYVRINNTSLLDIYQEDATFLESFLNTIGTDDTRSIJQDLMNRLTFISKV
CKGSGLNDYANYVRINNTSLLDIYQEDATFLESFLNTIGTDDTRSIJQDLMNRLTFISKV
CYANNTVIATDRSCSLIPHADHAIFLGPGSCPQGFLMDSDTEVCPSVDLHANVPOTEV
CPKAPLSISKANHTRGSDRTLKVNLSIHHIQNLKVSAPLHALVAIGGVSGSGKTSLLLEG
FKKQAELLIAKGTTTFSDLVVIDSHPIASSGRSDISTYFDIAPSLRAFYASLTQAKALNI
SSTMFSTNTKGQGCSDCQGLGYQWIDRAFYALEKRPCPTCSGFRIQPLAQEVLYEGKHFG
ELLHTPIETVALRFPFIKKIQKPLALLDIGLGYLPIGOKLSSLSVSEKTALKTAYFLYQ
TPETPTLELIBELFSSLDPIKKQHLPEKLRSLINSGHSVIYIDHDVKLLKSADYLIEIGP

TPETPTLFLIDELFSSLDPIKKOHLPEKLRSLINSGHSVIYIDHDVKLLKSADYLIEIGP GSGKQGGKLLFSGSPKDIYASKDSLLKKYICNEELDS 124549 126006

CPD_0097 124549 126006
pyk-Pyruvate Kinase
DSMITRTKI ICTIGPATNSPEMLAKLLDAGMNVARLNFSHGSHETHOQA IGFLKELREQK
RVPLAIMLDTKGPEIRLONIPQPISVSOCOKLRLVSSDIDGSAEGGVSLVPKGIFPFVPE
GADVLI IDDOYTHAVVVSSEADSLLEFMNSGLLKSHKSLGIRGVDVALPFMTEKDI ADLK
FGVEONMDVVAASFVRYGEDIETMRKCLADLGNPKMPIIAKIENRLGVENFSKIAKLADG
IMIARGDLGIELSVVEVPNLQKMMAKVSRETGHFCVTATQMLESMIRNVLPTRAEVSDIA
NAIYDGSSAVMLSGETASGAHPVAAVIKMRSVILETERNLGHDSFLKLIDDSIGALQVPSY
LSAIGLAGIQIAERADAKALIVYTESGCSPMFLSKYRPKFPIIAVTPSTSVYYRLALEMG
VYPMLTQESDRAVWRHQACIYGIEQGILGNYDRILVLSRGACMEETNNLTLTIVNDILTG
SEFPET

126091 CPn 9098 127494

CPn_0098 127494 126091
NO CODUST HOMOTOR PRESENT IN GENERALIK/EMBL AND OF TIT/7/98
LVOKKEHOTIKET LEAPLYYLVSOT TAL/EMITPRISED/DUCKEGGGLAFYT LODYRKTAL
TNLALAFPEKTFOBRYK TAROLOHLT ITLLEILIA LEOLUVIK IDKLITTUTEGRNPKGES
SEEVIGUEDLEETEKNIJGEKO IL LEFYSHQANWELBEFLY TEKNYIK FIAPAKAT KNORLEK
KIFALHEVEKOK IVPEKNITQOG TEALHYJEKLYTIVOTKOALLMISYTYPLEGGPAFTTE
PALLAYMOGEV TAVINGROAKGEEV I EDAKLYANKELIMPICIVAT LIMLOMMETELEKG TA
SOLFEJAMMIHERWIKKI KONVIKKYRYZCHILLYRIOOVEDHEDELRALABGES STITLHLAL
GNADHLEELOEOFFEYGLIOLKNOD ILALEKCYPA I EDLITINI JOHLYKHERKT GCAVY
CKEFLEKSLDHPOAPLKNELKIETYOPHILKIDKEPENEKVEGGERELIVE

127527 127805 177865 CPI_0099 (27727) A7800 (FFT)
No robust homolog present in Genebank/EMBL as of 1177798
YYCASYYLKISRFAAKHAPERIMERS/CYTTPTPTRATELLIPKVIFEARSTTPVQIKMISIK ELLTANGUES DADANALE YEDGEFÜGE ÖNNAN FLYVÜRE KAGELERTE ERLE KELFTERE.

127882 129133 CPn_0100 CT011 hyporhetical protein RTOKKTFILLDLETMIKFLSOLFIRHWPRKW/SLGFA!!!WILVGQSVT!TRTLTNVPVR RTOKKTFILLDLETMIKFLSOLFIRHWPRKV/SLOFAIIIWILVQGS/TITRIIN/PVR
IVOLHPDOTVLSIQKSGELNKKYSITITOHKNIT/ODLRPSNLEVVISAANHTESMIATIG
IVOLHPDOTVLSIQKSGELNKKYSITITOHKNIT/ODLRPSNLEVVISAANHTESMIATIG
KHNLVSVOHBINITKOIHSVDANDIFVRLTQY/TEDILLTITKPIGSPKKYETLDWPK
YLKOKVSGPKEYINALKEGGLELTFNLKKISFEELERNRIAQGSHDEIIFFIPKEMKKIL
IPPENTFMDLNDPOADFLRILFILKRECIPLINLIKEJFEETTETOTHNPLEYSLDPVPP
ISTREITHEISTETCYPLYJSP, FERAVYRDSRYJSTUMBSEDDDTGISMAIGELFEKTLEN
HELGETTAGESTILHDIALIDEN TEHBERSECURTALETAGESPLUKTAEIENDVVIO TKTKETTKLYKKEW

CPn_0101 129986 129141
ybbP family hypothetical protein
PSTLCNFSQYTTQGPSKTMPFDITYYTTPLLEIILIWWMLNYLLKFFWGTRAMDVVFGLL
AFLFLFVLADKLHLPIIRRLMLHVVNIAAIVVFIIFQPEIRLALSRIRFHGKKFFIDTQE
QFVEQLAASIYQLSERGIGGLVVLENKDSFDEYLSFSSVKINATFSEELLETIFEPSSPL
HDGAVILRGDILAYARVVLPLAHDTTQLSRSMGTRHRAALGASQRSDALIITVSEENGSV
SLSRDGLLTRGVKIDRFKAVLRSILSPKEHKRKPLFSWIWKR

CPn_0102 130099 131466
CydA-Cytochrome Oxidase Subunit I
FY1OFWKFMDALILSRIOFGLF ITFHYLF-VPLSMGLSMMLVIMEGLYLVTKKQIYKQMTW
FY4OFWKFMDALILSRIOFGLF ITFHYLF-VPLSMGLSMMLVIMEGLYLVTKKQIYKQMTW
FWVG1FALTFVLGVVTGIMQIFSFGSNMANFSEYTGNTFGTLLGSEGVFAFFLESGFLGI
LLEGRIKVSKKMHFFSTCNVALGAIMSAFWIICANSWMQTPSGYEMVMHKGKLIFALTSF
WGVVFSFTTIDRF IHAVLGTWLSGVFLVISVSAYYLWKKRHHEFAKQMKIGTICAVIVL
VLQLWSADVTARGVAKNQPAKLAAFEGIFKTEEYTP IWAFGYVDMEKERVIGLPIPGALS
FLVHRNIKTPVTGLDGIFRDEWPNVQAVFQLYHLMIMLWGMVALTLISWSAYKGWRWAL
KPFFLVILITFSVLLPEICNECGWCAAEMGROPWVVQGLLKTKDAVSPIVQANQIVQSLVI FSLVFIALLTLFITVLCKKIKHGPEEENDLTEFEVK

CPn_0103 131465 132511
CydB-Cytochrome Oxidase Subunit II
NRGIFMELSLTSLLPLAWYYLGVAVFAXSFCGGFDLGLGAVYLKAKEDKERRILLNSIG
PVMDGNEVWLVIIVGGLFAGFPACYATLLSIFYMPIWTLVLLYIFRGCSLEFRSKSESVS
WKIFWDIIFICSGTAISFFLGTIVGNLILGLELSPDTSYASLSWILFFRPYAALCGAVVA
SAFAINGSCFALMKTSDSLMARIAGOFPYILSSFLVFYVLFLGASLISIPKRFDAFPTYP
LILLILALTSCCCVAAKTSVSKKRYGYAFIYSTLNLLSLILSAATLTFPNILLSTVDPQY
SYTIYNSAVETKTLKSLLIIVLIGLPFIITYTCYIYRVFRGKTNFPSIY 131465

CPT:0104 133884 132676
CT017 hypothetical protein
EKSEMHLOISMLLIALGTAINSPAIYAADSQSVSFPEQLPSSFTGEIKGNHVRMRLAPHT
EKSEMHLOISMLLIALGTAINSPAIYAADSQSVSFPEQLPSSFTGEIKGNHVRMRLAPHT
DOTTIIRFFSKGDLVAVIGESKDYYVISAPPGITGYVFRSFVLDMVVBGEQVNVRLEPSTS
APTEVBLISGTQIQPASQEPHGKWLEVVLPSQCVFYVAKINFVANKGPIELYTQREGKKI
AMDLINSALNFAHIELEKSLNEIDLEAIYKKINLVQSEEFKDVPGIQGLIQKALEGIQKA
TÄÄKSLESQNTSIASSQCSTPKVSSSSVTTSLLSRHIRKOTALKTAPLTQGRENLEYSLF
RIMASMQQGNDHSEALTQEAFYRAEQKKKQVLAGVLEVYPHVVKNNPGDYLLKAQENTIA
FEYGTSINLEQWLGKRVTVECLPRPNNHFAFPAYYVVGIKEAS 132676

CPn_0105 134029 134883 CPn_0105 134883 134029
CT0166 hypothetical protein
VPFRKFSNONPMLLIYCKKKEIHLOWPOTAKIRFTPKIAMKVKINDQLICIPPFISARW
SOEFFIESOBGENKDGTLRIHLIDGKIISIPNLDQSIIDIAFQEHLLYLETSQSGKEDS
RDDDKLGVGVLMVILQQITKGNDIQVLPKNLISPLFSGTNPIEAILQHTPEHKDHPDAPT
DVALEKMADVIRVLSGNNATLLPRPEPHCNCMHCQIGRVMNEEDTLAVSDKDLTFRTWDIM
OSGDKLYIVTNPLNPSDQFSVYLGPPIGCTCGEPNCEHIKAVLYT

CPn_0106 phoH-ATPase 135073 136374 phöf-atpase
EKWETOMKKTMVIDTSVFIYDPEALFSFENTRIIIPFPVIEELEAFGKFRDESAKNASRA
LSNTRLLEMAKTKVTDGVLLPSGSELRIEVAPLSNDDRRGKLITLELLKIIAKREPMVF
VTKSLGRRVRAEALQIESRDYESKRFSFRSLYRGFRELQVSQEDIENFYKNGYLDLPLDV
VSSÖNEYFFMSAGENHFALGRYVSEGKIIALKANDKSVMGIKPLNTEQRCALDLLLRDD
VKJÖTLIGQAGSGKTILALAAANHKYPOKETYNKVLVSRFIVPMGRDIGFLPGLKEDKLM
HMÄÖPIYDMMEVLFSINQMGNSSEALQALMDAKKLEMEALTYIRGRSLPKAFIIIDEAQN
LTHEBIKTIISRAGKGTKIVLTGDPTQIDSLYFDENSNGLTYLVGKFHHLALYGHMFMTR TERSELAAAAATIL

136392 137321 CPH_0107 137321 136392
CT058 hypothetical protein_1
KKSPPPVTPKEIPTOPKPPIPORPEVSPTPTDHIVPGSIEASPILGKKPSPDSMVSPLSL
FHKMLLENMTPVEEPFFWPPAEKNOKIFAWALNOSKLIFVSTSGNIAOPRLVTDSMSMI
VNAANRTMSRDGAGTNOVLSAAVSVDSWGLSORPLNPERQGTPLMEGECRAGMWRNAGOS
NHTOKOGKPHYLAQLLGPKAVDHHNKSQAAFDRCKNAYLNCFSLAOTLGVTFLQIPYISS
GIYAPPENRKKPHSEENKVRMRWIHAVKCALVAAMQEFGNEPGNTDRRMLIVLTDLKTPA
TTDRKKKSHI. ITDPKKKSHL

137887 137303 CIVIO KNIFHYKAILMSIFNEEVFIISHRHTPLGGTSTALRNTPLVNPLHRTNLORIASYIPIFS TFIGIKTLKGISSLQYSMVLMTGNFSSVCKTLPCPEIYEELPKVRKEAWLEIFGIKALYY LVLGVIKIIKLIVRYLCPCCRPPEPREPONPLTPTPLDMGQQIDAIFSTPTJPTGFKDPF

CPn_0109 138646 141783
11es-Isoleucyl-trna Synthetase
RQMMTADEVCKNISFAKKEEQVLKFWKDNQIFEKSLQNRQGKTLYSFYDPPFATGLPHYG
HLLACTIKDVVGRYATMDGYYVPRRFGWDCHGVPVEYEVEKSLSLTARGATEDFGIASFN
EECEKIVFRYVHEMEYY INRIGRWVDFSSTWKTMDASFMESVWVFQSLYNVGGLYVCGTK
VVPFSTALGTPELSNFEASONYKEVDDPSLVVRMPLQNDSASLLWYTTPPWTLPSSMAIAV
GETLYYVRIQDKKGEQWILSQCSVGWFSINPEETVILEFSGKDLVGRTYEPPFTFFQS
KREEGAFRY LAASFVEESEGTGVVHMAPAFGEGDFLVCKENHVPLVCVEDAMSFFTEEIP
QYCGCYIKHADKEI HRVLKKEGRIFYHGTVKHRYPFCWRTDTPLYKAVNSWFVAVEKIK
DKMLPANSI HWYDEHIORARFCKWLEGARDWAIDRNRYWTTP UPINKSACSEILVVGSI
RELEELIFYTQ TTO HTHIE I DOLAH IVNCKRFFHIR EFVFFCWFDGAMPYAGHIYPFENOK
ETEEAFPADFIABGLDQTRGWFYTLTVIGAILFDRPAFRNAIVNGIILAECJIKMSKRIN
NYCSPKYVLDTYY ADALRIYLLBSVVVKASADLRFSDKGIESVLKQILLPLTHT/LSFFNTY
MELWEIPPENOX LEDPATEEDWONLSNIKSVVGKWFSMSJYHLMFAVEPFTFIDDLTN
WY LIMCHRREWEAEDTPDHRAAFSTLYEVLTVECKVIAPFVPFLAEDIYQKLKLEKEPES
VHLCGFTQVEMDKILPDLEKRMHDIREIVGLHERKEKKKRQPHANFY/JGSKORLS CPn_0109 138646 141783

LIKTFEGLIAEELNVKNVII FIYTT/KPNFRMLGKKVGSKMKEVQKALGEPUIVE DVYICBHTDMYTARSSALFSVILDCOLREPLIVE GIARELNKINTHRRNOQLH7/SDRIALRIKTTEAVHRAFLDYENYICEETLIIAYDFTQD SDFQGENWDINGHATQIETT/SSIDS

41827 CPn 0110

LEGGR LIMITED BLEEF OFFI SATING THE ALLASTIC BOLE TRADETRIBELISTS KRY LKRCMGREJDETY FIRM ALLOSAURI EFFEVYIGLESTYNYPY LEFOLITEJJITE GOKTI 10FKOFNOSYGRLI FPOTEMYGOFFDIKEWHODEPINKLKDPHLSFVSYADLIFGMG MYAMYR LLTEHQARTSHLLPNPGSPTKVYLEICHTANLSYPKPLLRHYEHOLSPAIOPMK YTALPLRKEHLHLI RINLITTSFF/LVAQGCAYKYHQFKINTSGI AKAYAILLPKVPDGCYEY SKGEAYQIGFGE IRYKLKSSHPLTQLNDKQVIELFNCGINFSSIYNPVNPLQAPLPNRYA FFNOGNLYIMDSPVFI KNDPYLQKFVTSETEKQEGSSETQPY IAFVDKGLPPEDFKEFVE FIHNFGIQVPKGHVLVLGLIMPMSADSREFGFVPMENLLGSPLCTFWPIGRMGRLTGVSA PTTLSGYLVSGIALATGLSYLGYVYYQKRRRLFPKKEEKNHKK

143934 44761 CPH_0111
CT021 hypothetical protein
CLONRYPIMPNDSSTYFERILOKYLMKKOGKTLFLFLFLSFLFSTAFSGLFASGTSSLRT
GLONRYPIMPNDSSTYFERILOKYLMKKOGKTLFLFLFLSFLFSTAFSGLFASGTSSLRT
IOENIFLAKTGDYTVLSGSGRFFYLVKSTTYFKTWHIEIHFPCIAHKERPSLEQASWKT
VIHQLESPSQVFVVSU/SSEGSGFFSLINTRTKSLEPVGKSTTVPAFLQIFDLPLSPAPANV
VIHQLESPSQVFVVSU/SSEGSGFFSLINTRTKSLEPVGKSTTVPAFLQIFDLPLSPAPANV
VIHQLESPSGVFVSU/SSEGSGFFSLINTRTKSLEPVGKSTTVPAFLQIFDLPLSPAPANV
VIHQLESPSGVFVSU/SSEGSGFFSLINTRTKSLEPVGKSTTVPAFLQIFDLPLSPAPANV
WSIETPKGTSIVRAVDIGHGATSPYVYSLPDSKTQ

CPn_0112 | 144743 145093 gatB-(Pet112) Glu tRNA Gln Amidotransferase (B Subunit) DSDFGVVNMKKNTHPEYRQVLFVDSSTGYKFVCGSTYQSEKTEVFEGKEYPVCYVSVSSS SHPFFTGSKKFVDAEGRVDKFLKRYSNVRQPAQQPQPEEDALPAAKGKKKVVTKKKK

CPn_0113 145329 146405

pfrA-Peptide Chain Releasing Factor (RF-1)

gfmkkkvAp/tinklaeveiktisnpeifsnskeysalskehsyllelknaydkilnlekvL

addkQala/ekdpehvywleedinnenkveleklnikilesilvepdepddulvimelragt

ggeaalbygdcvrnyhlyasskgmkyevlsasesblkgykevymgisgrgvkrildyea

gthrvorypetetogrvitsaitiavlpepseedtellinebolkiuffrasgagggrin

vtdsavgithlpfgvvvtcodersohkikdkamrilkarirdaemokrhneasamrsagv

gsgdrserirtynfsonrvtdhrigitlynldkvmegdldpittamvshayhqllehgn

CPH_0114 146371 14751
hemk/A/G specific methylase
wpPtsysnmeikkalQedtaYLDYYGVPLSDCEALYILMDLLEVSSRAKLFDLVGISET
wLMEYRKRLALRGQRCPTAYLMGAVSFLGLERVDSRVLIPRTETELLAEYIINYLLESS
EIØTFYDICCGSGCLGLAIKKSCPHVEVVLSDVCPQAVAVANENAKSGLDVKILLGDLS
AYTRPADAFVCNPPYLSFNEIHLDPEVRCYEPWKALVGGSTGLEFYQRIAQELPKIVT
SYGVGWLEIGSSQGESIKNIFSKHGIYGRLHQDLSGRDRIFFLEMDGRDPVSSGAYS

Pn_0115 147279 148622

fth-signal recognition particle GTPase
MINSLEGKLSSIFSFLVSSRRINEENISESIREVELALLDADVNYHVVKDFISKVKEKIL
GEETIKHVSPOOGFIRCHHEELVAFLSDGREEFTIOKTPSIILLCGLOGAGKTTTAAKLA
DVYLKNKKKKVLVVPCDLKRFAAVDOLKILVAOTKAEFYOSOENKPINVVVKALAYAKE
NGMOFVILDTAGRLNIDNELMEELTAIQKVSQANERLFYMNVAMGODVLATVQAFDQSLD
LKGVILSMTDGDARAGAVFSIKHVLGKPIKFEGCGERIQDLRSFDPOSMAERILCMGDTI
KFVKEMERY ISEEEDAELGKKLVTAAFTYEDTYKOMKAFRRMGPLRKLLGMMPGFNNAKP
SQKEIEDSEOOMKRTEAIILSMTPEERKELVELDMSRMKRIASGCGLTLGDVNQFRKOMS
OSKKFFKGMSKGKMEQVRKMSGGNOWR

Pn_0116 148592 148972
rs16-S16 Ribosomal Protein
ekwyrrksvalkirlrogornhuvyrlvladvesprogkyiellgwydphssinyolks
erifywlergolsskaealvkogapgvysallskoearklvvrkkrrayrorrstoree AAKDATK

CPn_0117 148983 150071
ETMD-tRNA (guanine N-1)-Methyltransferase
TGMKIDILSLFEGYFDGPLQTSILGRAIKQRLLDVQLITHLBOFGLGKWKQVDDTPFSGGG
MLLMAEPVTSAIRSVRKENSKVIYLSPQGALLTAEKSRELAAASHLILLGCHYBGIDERA
IESEVDEEISIGDYVLTNGGIAALVLIDAVSRFIBOVLGNOESAERDSLENGLLEGPQYT
RPREFEGKEVPEVLLGGDHKAISOWRLEGSERRTYERRPDLYLMYLYKRSIDHKFDEETT
NRDHFRCDKISVVLEVNKLKRAKNFYCKVFGLDAMSCENKFCLPHEGKTIFWLREVQAE
KKNIVTLSLSLDCACEEDFCYLLRWELFGGKLLEKQADEHAVWALAQDLDGHAWIFSWH
BWK

CPn_0118 150075 150464
rl19-L19 Ribosomal Protein KKENFRWYIMVNLLKELEZEQCRNDLPEFHVGDTIRLATKISEGGKERVQVFQGTVMARR CGGGGETVSLHRVAYGEGNEKSFLLNSPRIVSIEIVKRGKVARARLYYLRGKTGKAAKVK EFVGPRSSKK

150520 . 151164 CPn 0119 CPT_0119 150520 151164
rhB-Ribonuclease HII
LMNTSISEIQRFLSMIAFEKELVSEDFSVVAGIDEAGRGPLAGPVVASACILPKGKVFPG
VNDSKKLSFKORAOVRDALMODPEVCFGIGVISVERIDQVNILEATKEAMLQAISSLPIS
PDILLUDGLYLPHDIPCKNIIGODAKSASIANASILAKEHRDDLMLQLHRLYPEYGFDRH
KGYGTSLHVEAIRRYGPSN:HRKSFSPIKOMCAIV

151125 151778 omk-CMP Kinase gmk-GMP KINGSE BELFONKANVOYOMNKILVDSPFOPDHOKOCUKLFTIGAPAGVOKTTLVRMLEQEFSSAF AETICVTTKKPREGEVIVKOYHFVSHEEFORLLDRQALLEWYFLFGECYGTSMLEIERIW SLOKHAVAVIDIQQALFIRSRMPGVSIFIAPPGOEELERRLASRGSEEGSORKERLEHSL TELAAANQFUYVI INDULNÇAYRVLKS IF LAEEHRHII.

151769 CHOIST hypother to I protein
EH MIKKOKPTNEKLAKLESSPEDLYNYATKOAKTKIAKODVESSAVATETLYLLDREGT OPERTEE I VYTASTITVERKRSEHTNOPKKODSAYTWODVK

CPN_0132 152062 153723
metG-Methionyl-tRNA Synthetase
CKVMPQKVLITSALPYANGPLHFGHIAGVYLPADVYARE 152052 CKVMPOKVL (TGALPYANGPLHFGH LAGV/LPADV/ARP. GDDVLY (CGSDEFG LAI TLANADREGU) YOEYVDMYHKLHKDTF EKLGFALDFFSKTTTPFHAELVDOFYSGLKASGL IENR I SEQLYSEQEQRFLADRYVEGTCPRCGFDHARGDECQSCGADYEA I DL I GPKSKI S CVELVKKETEHGYFLLDRKDALLSF I QGCYLPDHYRKEVVDY I EHVRSRA I TROLSWGI EVPDFPCKVFYWFDAP I GY I SGTMEWAASQONPDEWKRFWLEDCVEYVQF I GKONLPH LAVVE PAMELJOKLDYKVDALVVGDLVGGFYLLEGRQFSKSEGNYVDMKFLSSYSLDKLRVVL AAFA ET GREGOREFE GEROLFFANT OF THE PERSON BEGANDLING ARE THE POLED TO THE POLED AND THE POLED TO THE POLED TO THE POLED THE POLED TO THE POLED THE POLED TO THE POLED THE POLE CDDVLY (CGSDEFGIAI LFCACYCOKLLALISYPIIPESAVAIWEMISPKSLENCNLDTMYARDLWKEEILDVINEE **FHI.KSPRLLFTTVE**

CPn_0123 155975 153774
recD-Exodeoxyribonuclease V (Alpha Subunit)
nSMERICGYLEQILVENNDSGDITAYIKIPNTTFILLIKGKLPQPLELGSPIQIYGVWSH
SPSNTKYPOIHSYDSPLLYEYRGVFHYLTSKLIKGIGPKIAEKIIEKFOEKTCYVLDITP
ERLSEVSGISETRCVSICKQLCEQKILRKTLLFLQEYNIPIHYGVRIFKKYQEKSIEKIC
EDPFLLAREMEGIGFKTADFIAMKLGVPRNSESRLCAGIQHSLEELQEEGHTCYPIELLI
DVVAKLLNQDVFDTPITLEEIDTQILNMOKRKLLHIQDISGTLHVWTRYLHLAEKTIVSD
LKRILFSSRIRSIDGEKIAHVEENLSIDLAEQOREAIKACFSEKLLITGGPGTGKST
ITQAILKIFEQVTHSKILAAPTGKAAKRWTEITQKHSVTIHALLQVDFKTKSFRKNHDNP
IDCDLIIVDESGMMDTHLLHHFLKALPDYTTLVFIGDIHQLPSVGFGNILKDLITSNKMT
VIRLNKIFRQVHDSGIVTNAHRVNEGLEPILVSETGRNDFLFFQKDDQEEALNHIIHLVT
KFVPQKYHIYPQDIQVLAPMKKGTLGIYNLNKALKHALMPKKANLHGRFGSYAVGDKVMQ
IRNNYNKEVFNGDIGYVSTINFEDKAVVVRMEGKHVGYSFSELDDLVLAYATSVHKYQGS
ESPCIIIPHTSHFMMLYRNLLYTAITRGKKLVILVGTKKAIAIATRNNRVQHRCTGLAE
VLKELDTKKNYADL VI.KELDTKKNYADL

CPn_0124 156575 158068

No robust homolog present in Genebank/EMBL as of 11/7/98

IRSKORTVAITLLU/LGILLIASGIIFLAVAIFGLSSAVALGIGCGMTALGTVLLITGLVL

LIRSEKLALEQVEIKOARTRVNNELDOLSGVYYTYENVLDNIKKRSYRDLGFYRQAQEEV

TNLEQDIEEIFLTLRIRNALDNEEFFMTHAKQCLAQVGESLFQDASIDEFINLAHLSEI

RQHLDINDPRWSMTKKVKGTVVRFIYVSTMYKQIIKSHFEKSDFGQLRMGLLNNYKTIEE

VLYQSFGKGYNRAALLSEKTRIHTSILHMEKDEDKHLNIKNECASALEHFKKFRTLFL

GLSEEDVIDFTGASGWDCSKLPRKEVPLDGGKKKLRFKRTFADEQVGMDRTTSLEHMTP

QEEDPLDRLMDQVEQEATSVLKODBYWKEIETSEAKFRSLPREDFEKGSJOISYIRDL

DHLSVWANOLSAAEDALIEVTDVQEHGNREMLKNIQQGLELIEDAVKATLPRVDFIQEL

LEKEELPLVAARMSLENS

CPH=0125 158072 158605

No.Tobust homolog present in Genebank/EMBL as of 11/7/98

KISSCAEIMSEVKPLFLKNDSFDLATORFONLINNLQEQAEIYNEYEEKNARVQNEIKEQ

KDEVKRCIEDFEARGLGVLKEELASLTRDFHDKAKAETSMLIECPCIGFYYSIHOEEQRQ

ROERICKMAERYRDCKQVLEAVQVEQKDMISSRVVVDDSYFEEEKEEQKVDNRKKEQD

CPL-9126 158806 161085
Noi-EDBUST homolog present in Genebank/EMBL as of 11/7/98
Noi-EDBUST homolog present in Genebank/EMBL as of 11/7/98
Noi-EDBUST homolog present in Genebank/EMBL as of 11/7/98
LLVFSYYCMGLFFFSCAISSCGLLVSLGVIGLISVLGVLLLLLAGLLLFKTQSMLREVPK
APDLIDLEDASERLRVAKASFSLASLPKEISOLESVIRSAANDLNTIKTWHKKORLVFSV
SREERLAAAQNYMISELCEISEILEEEEHHLILAQESLEWIGKSLFSTFLDMESFIMLS
HLSSVRPYLAVNDPRLLEITEESWEVVSHFINVTSAFKKAQILFKONEHSRMKKKLESVO
ELLEFFIYKSLKRSYRELGCLSEKMRI IHDNELFFWVQDQQKYAHAKNEFGEIARCLEEF
EKTFFWLDEECAISYMDCMDFLNESIQNKKSRVDRDYISTKKIALKDRARTYAKVLLDEN
PTTEGKIDLQDAQRAFERQSQEFYTLEHTETKVRLEALQQCFSDLREATNYNGVFTINSE
NANDLKESFFKIDKERVRYQKEQRLYWETIDRNCQELREEIGESLRLQNRRKGYRAGYD
GREKGLLRQWKKNLRDVEAHLEDATMDFEHEVSKSELCSVRARLEVLEELMDMSPKVAX
LEELLSVEERCILDIFRENLERAYLQVMKCSEILSKAKFFFFEDEDCLLVSEANLREVGAQL
KQVGKCQERAQKFAIFEKHIQGQKSLIKEQVGSPLAGVGFLKSELLSIACNLYIKAVV
KESIPVDVPCMOLYYSYYEDNEAVVRNLLAMTERYQNFKRSLNSIQFNGDVLLRDPYYQ
PEGHETRIKERELQETTLSCKKLKVAQDRLSERSR PEGHETRLKERELQETTLSCKKLKVAQDRLSELESRLSRR

162152 161130 VILETCATIONIC AMINO ACID TRANSPORTER
ESEMPESANCESRERNVPLGIFHGLVACLYWGIVFVIPNFLGSFGDLDIVLTE ESFMFFSANQESRTRNVPLGIFHGLVACLYWGIVVVIENFLSEGDLUVLIWITETSIF SLLÄGAIKHPSVIKKTPLVIMRKSLLUMTLLIMPVYFFGITLGIRVVGSAITVIISLAPT AVLYHSNTKQKELPYSLLFAISSVIITGVILTHLSALNLPTAASPLYSILGVIAVILSTS LWVIYVIRNQSLLEKHPNLTPDTWSYLIGISALIICLPMIIILDLCGITHYTHNLISHTP GSERLLFLLLCSAMGIFSSAKALIAMNRASLNLSPALLGAILIFEPIFGLVLTYLYSQSL PSLQEGIGIFLMLGGSLLCLVLFGRKVQKSLENSQVSSSNE

162262 163053 CPH_0128 163282 163033 bpl1-Biotin Protein Ligase EDRGRMLRNQVLVYCSEGVSPYYLRHTIRFLKYYSTOEGAFDILWDGNFLIKNPFWEET TRLLVPFGGADRPYHRVLHGLGTARIFQYVSEGGNFLGICAGAYFGSKMIYFYEPEGAPL GGARDLGFFGTAKGPAYRGNFSYVSPSGVRVSPQLFSDFGLGYAMFNGGCFFEGSEGYP GVNLESRYDDLPGKPASIVSRJVSKGLAVLSGPHIEYLPHYYRMVKENVQKTREFLQRER TTLDRYCONLVORLROPAFSKADC

CPn 0129 163747 163064 CPD_012?
similarity to CT036
pegyilshihmdprifytseploktygokloekhvanlilasovsltdlonktoyennlie
ttyneltyyepvyhnpdilrsemdpisnolylifkkfihyhnlfstalenngillidsl
htgssnelaromellafloypeoldynedeytieppvyhnprvyknsotapoiosfgllh
gyeemsyasnnirnvlthsivlcspilyolitefptykhhaddfdcli

CPN_0130 164251 163751
No robust homolog present in Genebank/EMBL as of 11/7/98
SSHWKCSSTENENKRPAGLEPESKRPAATRKLYLATLGEFLGTAACILIALSGLLENKILI
TALGLIGI INJERGIGLI LOTOGGSKOVOKDEØKREN ITFRKENPGLEPPGLPPULNPLKNKIQSS
ETILLDPYSINLKNELFFPDFEEWKKIFLKPPDFLIKSALANWKILE

No robust homotosi present in Genebank/EMBL as of 11/7/98 essekkeresistalfilifeftsavves icelelemenamessevynopowilktsvaq evekkeresistalfilifetsymletelisyvalteteytytalmlamstvileliksv RESMODERM TERRITATIONERS PLANCES DE LE LETTET ENVELMES DE L'EDENSOAA VELLE CHAPTER DE LE CHAPTER DE L'EL ALLKNPOGISIKDLKOFLVH

CPN_0132 165584 166561
No robust homolog present in Genebank/EMBL as of 11/7/98
SMIEFAFVPHTSVTADRIEDRMACRMNKLSTLAFTSLCVLISSVCIMIGILCISGTVGTV
AFVVGIIFSVLALVACVFFLYFFYFSSEEFKCASSOEFFFLPIPAVVSALRSYEYISQDA
INDVIKDTMQLSTLSSLLDPEAFFLEFPYFNGLIVNHSNKEADRLSREAFLILLGEITWK
DCETKILEWLKDPNTTPDDFWKLLKDHFDLKOFKKRIATWIRKAYPEIRLPKKHCLDKSI
TKKTCCFETLILEBETYCKTSTLBKLTYFEGGFFBANTIGIGGEVFMMIGLPKVPVDLTWEMF
MEIMHOULD GFFFBBART LBLVALL

CPn_0133 167349 66564
CHLPS hypothetical protein
NSSAYMFKLLKNLFLIGGCIVGYFMARKESIVEOWLSNRLHTQVTVGRVSIRTSGIKIRH
ICIHNPLASERFPYAAEIEYADVRSSISMLLTKOLEISELIIHGANFTIFPYDSHGTKT
WWSLVWKNFHPOKETPSNLWIDRAEVLIRRCLFLNTRLYGLRANHKDIPHLSVPSLEFHS
HTSSAKELPKLSEALPSLLYLALEESLYHLALPGDIIKPLSQAHKHFYSSYPOFODRLN
DINTPGTPTEEIIGFIRGLFFH 167349 66564

167467 CPn_0134 groEL-HSP-60 169131

GTOEL-HSP-60

PADYRKLRYTTMAAKNIK/NEEARKKIHKGVKTLAEAVKVTLGPKGRHVVIDKSPGSPQV
TKDGVTVAKEIELEDKHENMGAQMVKEVASKTADKAGDGTTTATVLAEAIYSEGLRRVTA
GANPMDLRRGIDKAVKVVDELKKISKEVOHHKEIAQVATISANDISEIGNLIAEAMEKV
GKNGSITVEEARGFET/LDVVEGMFNRGYLSSYFSTNPETQECVLEDALILIYDKKISG
IKDFLPVLQQVAESGRPLLIIAEEIEGEALATLVVNRLRAGFRVCAVKARGFGDRRKAML
EDIAILTGGQLVSEFLGKKLENTTLAMLGKAKKVIVTKEDTTIVEGLGNKPDIQARCDNI
KKQIEDSTSDVDKFKLGERLAKISGGVAVIRVGAATEIEMKEKKDRVDDAQHATIAAVEE
GILPGGGTALVROTPTLEAFLPMLANEDEAIGTRIILKALTAPLKQIASNAGKEGAIICQ
QVLARSANEGYTALRDAYTDMIDAGILDPTKVTRSALESAASIAGLLLTTEALIADIPEE
KSSSAPAMPSACMDY

169448 169143 CPn_0135

groes-10/Da Chaperonin MSDQATTYRIKPLGDRILVKREEEEATARGGIILPDTAKKKQDRAEVLVLGTGKRTDDGT LLPFEVOVGDIILMDKYAGQEITIDDEEYVILQSSEIMAVLK

CPn_0/136 171419 169569 CPL_0/36 171419 169569
pepF/01igopeptidase
KGVPSLMTTELKTEALPTRTQVDPKHCWDTTLMYANREEWKKDFDLCSSGKDRSPIWPEF
SP\$MYQIDNPESLLELLSKKFSVERKLDQLYIYAHLIHDQDITNPEGESDYQSIVYLYTL
F90EISWIQPALIALSEEKVAALLSSSVLAPYRFYLEKIFRLSPHTGTANDEKILASSF
MLNYSNKAFSSLSDAEIPFGIAKDSNGEEHPLSHALASLYMQSPDQELRRTAYLAQFQRY
/DYRNTFANLINGKVQAHLFEAKARNYPSCLEASLFQRNIPTTVYINLINETKKHTSLIN
RYFNLKKEALNLKEFHFYDVYAPISQTTSKNYSYEEGVDLVCKSLLPLGTHYVEILRING
LSNRWVDRYENKHRRSGAYSSGCYDSAPYILLNYTNTLYDVSVIAHEAGHSMISYFSREA
QPYHDAQYPLFLAEIASTFNEMLLMEALSKSDQSKEDKIVIITKTLDTIFATLFRQTFFA
AFEVEIHSAAEQGTPLTEEFLSATYGNLQKEFYGGVYTSDSUSSALEWARIPHFYYNFYVY
QYATGIIAALSFAEKILTQEPGALELYLKFLKSGRSDFPLNILKKSGLDMTTSAPLDKAF
AFITKKIDLLSSLLSED

172263 171502 CPn_0137 CPH_U13/
ybg1-ACR family
vCSMNVADLLSHLETLLSSKIFQDYGPNGLQVGDPQTPVKKIAVAVTADLETIKQAVAAE
ANVLIVHHGIFWKGMPYPITGMIHKRIQLLIEHNIQLIAYHLPLDAHPTLGNNWRVALDL
NWHDLKPPGSSLPYLGVQGSFSPIDIDSFIDLLSQYYQAPLKGSALGGPSRVSSAALISG
GAYRELSSAATSQVDCFITGNFDEPAWSTALESNINFLAFGHTATEKVGPKSLAEHLKSE **FPISTTFIDTANPF**

174094 172700 CPn 0138 CPn_0138 174094 172700

hemL-Glutamate-l-semialdehyde-2, l-aminomutase
TMSRLFILAIKDOLLONMMKLTKRISMLNCSNOKHTVTFEEACQVFPGGVNSPVRACRSVG
VTPPIVSSAQGDIFLDTHGREFIDFCGGWGALIHGHSHEKIVKAIQKTALKGTSYGLTSE
EEILFATMLLSSLKLKEKIRFVSSGTEATMTAVRLARGITNRSIIKFIGGYHGHADIL
LGGISTTEETIDNLTSLIHTPSPHSLLISLPYNNSQILHHVMEALGPQVAGIIFEPICAN
MGIVLPKAEFLDDIIELCKRFGSLSIMDEVVTGFRVAFGAQDIFTLSPDITIYGKILGG
GLPAAALVGHRSILDHLMPECTIFOAGTMSGNFLAMATGHAAIQLGQSEGFYDHLSQLE
LFYSPIEEEIRSQGFPVSLVHQGTMFSLFFTESAPTNFDEAKNSDVEKFQTFYSEVFDNG
VYLSPSPLEANFISSAHTEENLTYAQNIIIDSLIKIFDSSAQRFF

174686 174093

yqge SPTKNKLRDIMKIPYARLEKGSLLVASPDINQGVFARSVILLCEHSLNGSFGLILNKTLG FEISDDIFTFEKVSNHNIRFCMGGPLQANOMMLLHSCSEIPEOTLEICPSVYLGGDLPFL OEIASSESGPEINLCFGYSGWQAGQLEKEFLSNDWFLAPGNKDYVFYSEPEDLWALVLKD LGGKYASLSTVPDNLLLN

CPn_0140 yqdE 175140 174673

yddi Prsnogkifomslekelleetpluvlnfyklusfonyagmilgteekkfaiyghusmoga Fogadteghsporpfahdllnfufsgfdigvlrvuindykdnufytrlfleokdreflyu VDVDARPSDSIPLALTHKIPILCVKSUFDAVUPYEE

175817 175110

CPN_0141 1/381/ 1/5110
rpia-Ribose-5-P Isomerase A
HSSSAVENDLHLHEKKCLAHEAATOVTSGMILGLGSGSTAKEFIFALAHRIQTESLAVHA
HASSAVENDLHLHEKKCLAHEAATOVTSGMILGLGSGSTAKEFIFALAHRIQTESLAVHA
HASSOMSYALAKOLAIPLLNPEKFSSLDLTVDGADEVDPQLRMIKGGGGAIFREKILLRA
AKRSIIL/DESKLVPVLGKFRVPLEISFPGRSAIIEEIRHLGYEGEWRLODTGDLFITDS SNYTYDTESPNSYPNPEKDLLKLTQTHGVTEVGFVTEKVEVWSSNSQGLTSKKYSV

176121 175816 No robust homotox present in Genebank/EMBL us of 11/7/98 SHSYSYSYSTEKFHEKILOLLSTKNSTLNFSSHFEISRVSHDNAIQKIRSYPLKPIAEN RINTLOFFICEKIDYPKDSSKREPFLYSIGPIPLHYLWEYFYT

177347 176214 **yxjG_Es_1 Hypothetical Protein
PRETMITTSLKRELKSHEDVVGSELREEHLKKTRESCKEGSISLDOLMGIEDIAIQDLIKK GAKKATA PENADYERAKTAN INDERACCINAS I DEKONDO TODATT INDEATOR OF HIBAADILA LADIES LA LADIE RPDDLVVNLIVCRGNYHCKFFACGCYDFIAKPLFEQTN COEKTVCLGLVTSKTPTLENKDEVIARIHQAADYLPLE EOWAKVALVKE [SEEVWK

efdhersgdfspltfi Ocgfasceignkltee

[77942 180560

CPD_0144 TYPE BUDDO
C1pB-01p Proteise ATPASS
KYLGVNPMEKFSDAVSEALEKAFELAKSSKHTYYTENHULLALLENTESLFYLVIKDIHG
KYLGVNPMEKFSDAVSEALEKAFELAKSSKHTYYTENHULLALLENTESLFYLVIKDIHG
MPGLAFTAVKDALSREPTVEGEVDEKPSPGLOTULRDAKQEAKTLGDEYISGDHLLLAF
ALDEGGERUMF, JTTYZGEYFFERBLUTT FREGREM, STEALEGRE, GERYYKKBLTALARE
GREEF VY BREEFFERPT FOALEMETTHIN MLGELST, SYTATARE GALFFT GVPPESLKET
CALFOR VY BREEFFERF TOALEMETTHIN MLGELST, SYTATARE GALFFT GVPPESLKET
CALFORNIA FOALEMETTHIN MLGELST, SYTATARE GALFFT GVPPESLKET
CALFORNIA FOALEMETTHIN MLGELST, SYTATARE GALFFT GVPPESLKET

KQLYVLDMGALIAGAKYROEFEERLKSVLKDVESGOGEHIIIIDEVITUVEGALK KQLYVLDMGALIAGAKYROEFEERLKSVLKDVESGOGEHIIIIDEVITUVGAGATDGAMD AANLLKPALARGTLHCIGATTLNEYQKYIEKDAALERRPOPIFVTEPSLEDAVFILRGIA EKYEIFROVRITEGALANAVLLSYRVIPDRFLPDKAIDLIDEAASLIRMQIGSLPLPIDE KERELAALIVKQEAIKREQSPSYGEEADAMQKSIDALREELASLRIGMDEKKLISGLKE KKNSLESMKFSEEBAERVADYNRVAELRYSLIPQLEEIKQDEASLNORDNRILGEVDE KERLIAQVVANNTGIPVORMLEGEAEKLLILEESLEERVVGQPFAVSAVSDSIRAARVGLND PQRPLGVFLFLGPTGVCKTELAKALADLLFINKEEAMVRFDMSEYMEKKSISKLIGSSRGY VGYEEGGSLSEALRRPYSVVLFDEIEKADKEVLNILLQVFDDGILTDGKKRKVNCKNAL FIMTSNIGSPELADYCSKKGSELTKEAILSVVSPVLKRYLSPEFMNRIDEILPFVPLTKE DIVKIVGIQMRRIAQRLKARRINLSNDDSVILFLSEQGYDSAFGARPLKRLIQQKVVILL SKALLKGDIKPDTSIELTMAKEVLVFKKVETPS

180717 182369

CPn_0145 180717 182369
CT114 hypothetical protein
NCASFIWLNKSSHRNLRSPMFKSFTVRYMFVGGLVSFLLPIPDLECANNVTKTYDKKAS
VISRDLKLQEDCQKFWNLDPYKLESLCAYQVLYHDDYSSKRIRELFPQIQKDEVPIFATM
ILTIGKVDRGFSPEEISLIQKLSYPGLSLASLRGSTEIDPNTDLARALVVSEFSGDLGKN
RADTYSNCLDILALRIHAERQRYLDQSPC-VPGTSEFHKATIEAINTILFYEEAVRYPSKK
EMFSDEFSFLSSVTDRKFGVCLGVSSLYFSLSQRLDLPLEAVTPPGHTYLRYQGGEVNIE
TTAGGRHLPTASVCDCLDLEDLQVRTPEEMIGLTFMNQGSFALQKKKYKEAEAFAYKKAQE
YLGDEELQEGLIGFVQILGGKKKEGKSLIGKSPRASQKGSVAYDYLKGRINIPTLALLFSY
PGSNYEEIASYEEELKKAMKSSMPCCEGGRRLASVAFHLGKTAEAVALLEKCVEDIPNDL
SLHLRLCKILCDRHEYTKALKYFIIAERLMEDQGFLKKDNRSFALFYEVKKIISKVAPQK
ANTLLLMESER ANTILLIMESER

CPn_0146 182595 183095
No robust homolog present in Genebank/EMBL as of 11/7/98
IIVGISMSSSEVVFOTVHGLGFGGLSSKSVVPFKKSLSDAPRVVCSILVLTLGLGALVCG
IATTCWCVPGVILMGGICAIVLGAISLALSLFWLWGLFSNCCGSKRVLPGBGLLRDKLLD
GGPSRAAPSGMGLPGDGSPRASTPSCLEELQAEIQAVTQAIDQMSDD

CPG_0147 183213 183671
No Fobust homolog present in Genebank/EMBL as of 11/7/98
HGGEMAVQSIKEAVTSAATSVGCVNGSREAIPAFNTEERATSIARSVIAAIIAVVAISLL
GLETVVLAGCCPLGWAAGAITMLLGVALLAWAILITLRLLNIPKAEIPSPGNNGEPNERN SATEPLECGVAGEAGRGGGSPLTQLDLNSGAGS

CET 0148 183822 185702
pkn1-S/T Protein Kinase
GRAWRYSSMESEKDIGAKFLGDYRILYRKGQSLWSEDLLAEHRFIKKRYLIRLLLPDLGS
SQEFMEAFHDVVVKLAKLNHPGILSIENVSESEGRCFLVTQEQDIPILSLTQYLKSIPRK
LTEDEIVDIVSQLASLLDYVHSEGLAQEEWNLDSVYIHILNGVPKVILPDLGFASLIKER
ILDGFISDEENRESKIKERVLLHTSEGKQGREDTYAFGAITTYLLFGFLPQGIPFMPSKV
FSDFIYDWDFLISSCLSCFWEFRAKELFPLIRKKTLGEELQNVVINCIESSLEXEVPDPLE
SSONLPQAVLKVGETKVSHQQKESAEHLEFVLVEACSIDEAMDTAIESESSSGVEEGYS
LATQSLLVREPVVSRYVEAEKEEPKPQPILTEMVLIEGGEFSRSVEGQRELEVHKVIL
HSEELDHPVTNEQFIRYLECCSEGDKYYNELIRLRDSRIGRRSGRVIEDGYAKHFVV
GVTWYGASGYAEWIGKRLPTEAEWEIAASGGVAALRYPCGEEIEKSRANFFTADTTTVMS
YPRAPPYGLYDMAGNVYEWCOGVYGYDFFEISAQEPESPQGPAQGVYRVLRGGCWKSLKDD
LRCAHRRINNFGAVNSTYGFRCAWIN LRCAHRHRNNPGAVNSTYGFRCAKNIN

187700 185706 CPn_0149

CFD_0149 185706 187700
dridT-DNA Ligase
ERFMKENSQAHYLALCRELEDHDYSYYVLHRPRISDYEYDMKLRKLLEIERSHPEWKVL
WSFBTRLGDRPSGTFSVVSHKEPMLSIANSYSKEELSEFFSRVEKSLGTSFRYTVELKID
GLAYAIRYEDRVLVQALSKONCKOGEDITSNIRTIRSLPHLIPEDAPEFIEVRGEVFFSY
STFGIINEKQQQLEKTIFANPRNAAGGTLKLLSPQEVAKRKLEISIYMLIAAGDNDSHYE
NLQRCLEWGFPVSGKPRLCSTPEEVISVLKTIETERASLPMEIDGAVIKVDSLASQRVLG
ATGKHYRWALAYKYAPEEABTLLEDILVQVGRTGVLTVVAKLTPVLLSGSLVSRASLYNE
DEHRKDIRIGDTVCVAKGGEVIPKVVRVCREKRPEGSEVMMPEFCPVCHSHVVREEDR
VSVRCVNPECVAGAIEKIRFFVGRGALNIDHLGVKVITKLFELGLVHTCADLFQLTTEDL
MQIFGIRERSARNILESIEDAKHVDLDRFLVALGIPLIGIGVATVLAGHFETLDRVISAT
FEELLSLEGIGEKVAHAIAEYFSDSTHLMEIKKMQDLGVCISPYHKSGSTCFGKAFVITG
TLEGMSRLDAETAIRNCGGKVGSSVSKQTDVVVMGNNPGSKLEKARKLGVSILDQEAFTN
LIHLE

CPN_0150 187759 192444

CT147 hypothetical protein
II'YKFFYSYNCPYFISFFVLLGVMMASSSNNSTKQDGIPSWVNPNVQMNRASQVGDEA
MSLTPEAQTSRSWFSDRKHFLEVLDVSLEEMENNDLKKYSRYKTIILIATLVTVAITGUP
PISMYFGIPMWPCLILFOAGLSSAFLSHRLQSKCKEHLHRYRAYQIYRQQLLSQYPDLR
KSTLYKYSITHVKPKKGFVCKLVENLRPDLHKNKDDCGAAADSRLDFAGYGVKHYYTDAL
LGVSGVNSVEWQRLASLIMSVKNDILNDVGSREPIDPAQRSALVVSGKDIGGEIDPGGIL
UGVSGVNSVEWQRLASLIMSVKNDILNDVGSREPIDPAQRSALVVSGKDIGGEIDPGGIL
UGVSGVNSVEWQRLASLIMSVKNDILNDVGSREPIDPAQRSALVVSGKDIGGEIDPGGIL
UGVSGVNSVEWQRLASLIMSVKRDLLNDVGSFRYDLASALVVSGKDIGGEIDPGGIL
GYSAKKFODLSALTTAHGTGGALEDLDSLLCYYDQLIESKGVGEKIIASIHOKHUDLAMQD
PCDQEHLKKWSNLYHVPSITIKEFTEGKLEQNEVVSRIQHLEGKLESKCSILINCKTNNA
EYATKSEKKLADYLLQIGDREPFLTGMHKAIATGKAIQGKVEGVISQHPEKOIMMLRGSI
EPLLEGMLRREDWIAILQKNEDEVLALKSTMEAQLOFFLDLVGTWEGKYGEFKNKLSKVL
VYOFTKINSYNLLNRLEVLHAESSTHODLVLHVDRMSEDLKKTIEEIDONLFGYTFEELSLL
AFEYQLIMNELPDLIVOEGNRLQEAISSEOVSGILMLIIGLINRDEKINKNYESSRKNLVA
IAKQAR:DARNID:UGLAPLIQRNRASLDNILQMYZLFNGSIRNIHALDTETLVATSSNM LAZGAR: DAKN LO IGOTLAPLI ORNRAS LONI LQNMY LENG SIRN HALDYETLVAT SSIM PSAMINT POWN LYTINLLOV LEIGSK PAPAPMEN POLLY ALPEEV QOAVAE DVSCTHRLIHQ VLEHRICADI.KIM BIQUAS INKWAMAKA LVLG IVAYLE CVUISA LE LOQUI LUSLLILISC VO LLLEDOVITAL FEDRI GISCISERE EKOVLETAOSIL LEATH I LEGERININD LA BLAKLQON LINLE GPGITTMARN LVIDLEGI TYKEKOLKOLTKEFRKOLYILINKRIKREK FIKTOLIQEA PVVR PT LPOLI IR INEVPARLIRE LEBILOKOKEETI I RODAL VJEDMOLCLEK JAYDMEKAHAAAMT KRIVIKLON LDRI JAKNMETY VRI LONFRITLIGEKLOF LTVOETOVIKEAKELHELAA LIYO NICHRICKOKAKKOPKEN VLHI MOKOQLELLEAYLIYTA OG ILCRIJOMQAS FRERI LLINE LOAMRIGEAERTI LITREMLKITUGLOYLITE FVRESIDE STOTISTANJULKVIR OLD LEOLI OMGETVI HEDYAAVQAALAAYVIKHEGLI VSTYGUSADE KOTSISTATIM KOLHAVEELV EMINIETYRLING DOLLIN VIKOVLIBILIDODSSONI I LDVVKKI FELLINNINNIN PNDPEC OKYMOILLDAPVSLLYGAFI FILMITEUNIAMSTKAAEEEAKRYVEEKGRGETY WEEAKORLEAIAAELDDLRI GIQEOYAEMQGIEDLELKOKFEDLOKKLEALEELLQIGRRIDSSVDKQKELLGLLGREE

625 194178 1.9 CPn_0151

CPI_0151
mhpA-Monooxygenase
CYENLEHYPRASMADILVIGANPTGLILANMLIOHGISVKVIDHRASPEDPSFLDCRKLP
CYENLEHYPRASMADILVIGANPTGLILANMLIOHGISVKVIDHRASPEDPSFLDCRKLP
CHILDER ELIMIGENT TOE FOANIE TO ARVEWYRPTGLEKEODATDOPVESSIGTTYO
MINGHLIDER ERFORD TOE MATERIALDIAND TOE TOTAL CHERIPRE TYDERM LAJEADN
MINGH WYGGLRARKINKE/IFING/EUEPFEEDHIHLD TTRIFLIRVYYNDJEKTKG HERNHLIDER FOR THE THE TOTAL AND THE SECOND TRUNCHER OF THE SECOND STREET OF THE SECOND SECON

CPn_0152 1952/4 194318
CT149 hypothetical protein
LINGRKVAFLVSCLFSVAIGASAAPVRVPGFPOIPEDLVQIKTEVCPKQEVCLAVTIKCD
DHNLIGVLHLPHIPPTEGGFFTVVLFHGFRGTKFCGLIGAYRKLGRKFAAVGIATLRVDM
AGCGDSGVAEEVPIETYLBDAOTILETVQEHPDLNAVRLGISGFSLOCHTAFELAKIYN
PROLNIKALSVMAPIADCG/LLKELYENFSKHGEGDIISVGKDFGFGPPPIIVCSGDVDL
LIRIODHVTANSLPTKPYJHOQGIDDTLVSRTQCTLFKNTAPGRMTFISYPNTGHNLAT

APDLDMILDQIVSHFQRTA

195430 197892 AKEAPKETVLSLSRSVVAKYLENAQIRKEIYVPNKLVNFVL

197874 199202

CPI_0154

97874

197874

199202

9SeA-MOO Transferase

TSEFCMMLRGVHRIFKCFYDVVLVCAFVIALPKLLYKMLVYGKYKKSLAVRFGLKKPHV
PGEDFLWWFHGASVGEVRLLEVLEKFCEEFFGWRCLVTSCTELGVQVASQVFIPMGATV
SILBELDSI IIKSVVAKLRPSLVVFSEGDCWLNFIEEAKRIGATTLVINGRISIDSSKRF
KFLKRLGKNYFSPVDGFLLODEVQKQRFLSLGIPEHKLQVTGNIKTYVAAQTALHLEREHT
KWIRLRLPTDSKLVILGSMHRSDAGKWLPVVQKLIKEGVSVLWVPRHVEKKRDVESELRH
LHFPYGLWSRGANFSVYPVVVVDETGLLKQLYVAGDLAFVGGTFDPKIGGHNLLEPLQCE
WPLIFGPHITSOSELAQRLLLSGAGLCLDEIEPIIDTVSFLLNNQEVREAYVQKGKVFVK ETASFORTWRALKSYIPLYKNS

199697 199488 CPn_)155 No pobust homolog present in Genebank/EMBL as of 11/7/98 NSISTGYPFLEKLKISLIPIEEMRHELFMKTHNSSSNGFSNQEKGIRTYFKSDLLGYEDL YPLRENINPN

CPn_0156 200147 199770 No robust homolog present in Genebank/EMBL as of 11/7/98 LGKOKLLARHMHNIVVLSEEFGRSAFLGRTAFFPNKYPIACGGVGIPSTIGNLFTIWYC FYFYRAATPQSDHPDGCGFILLERLKELGAGFFYCDLRESNTTGFTLFFEGSNKGVLKNH

CPn_0157 200753 200298
No robust homolog present in Genebank/EMBL as of 11/7/98
FSFVTYKEALMNIYOFSPGASFNWOASLMAQLNSYFCLGGETVTRIISLRPSGLILAKKE
KAVVSTAEKILKILSFILFPLVLIALAIRYLLYNKFNKDLDRAVFFIPTEITKAEELIIA KNPALVKEAALTVSPLFYSLPKKYQLMKVETP

200894 201463 CPn 0158 CPT_01D8 201403 200394
No robust homolog present in Genebank/EMBL as of 11/7/98
PPKITLSINIDLLLEDLOTDSIPWRKLYLSEDFDFAYYPESKAIIDTVAKLEKNNPGEEF
CLESKKILARYLLEQLFKLETGLNFPTSTIDGGRESFLIEFSHETKKPTWAFIVFYYH SNGPKLEKDFKQAGCEVHNRLLNLGLKYRPQAGAQNDGRNGGPYGP IGFLIVWEENYGSV LKDHGFIKDN

CPn_0159 201811 201467
No robust homolog present in Genebank/EMBL as of 11/7/98
CCPGGETATRIFSMTPSGFSLATEEK/VOYSTAEKVIKILALIFFFIILIALAIRYFLHRK
FDRKCFVIPCDTPKELELILAANPQLVEKAAREVHPGFFALPTKYQSMYIQTSKG

CPL_0160 203794 202127
pfkA-Fructose-6-P Phosphotransferase
TVELLSLNKSYFEIGRLRYRPEILTLLETIRSKHIQETSSPPSPPPELQKHIPNLCRIPE
VSIYTEOFFSSKPLKIGVLLSGGQAFGHNVVIGLFDALRVFNPKTRLFGFIKGPLGLTR
GLYKDLDISVIYDYYNMCGFDMLSSSPEKIKTEEQKKNILNTVKQLKLDGLLIIGGNNSN CLYROLDISYTYNYNMAS PURSSERENTRI EDYRANIA TAN QABAGETT SENTGALASY TEAHNARTS LOVENS TO TOTONLAS TEAHNARTS LOVENT TO TOTONLAS TEAST TEAHNAST TEAHNAS TEAHNAS LOVENT TEAHNAS TEAHN TPVTRTDSVDPKSPAVQHULQQCDCCL/EDLYRFPCIPLQYFCKEELTDQRPLTLLWENQTHSPFQALYSTSGKRSL

CPn_0161 204658 203798 (predicted acyltransferace tamily) HROSTORKGEPEL FOLVVLERHEDFTMESETELINFTTFGLERTFLIFYNFFYP I VILLHG LACIMTOSKROHVRLAGELTRIG I AALRVOLLGHODG BELMDFOLENY KOM I REI I EYT HOLLHI DQERLA I FOOSLKYTLAID/FLIFFNK I KALAVWALT I GRELMAAFAGKNAPEV I TM::QKGAITYAGMTLNPDFYTQFLKIDIVKELMPSARN FTEAFANQDKPITILTYPDVDHAFPFAESSALSDLTQI

YMQGEQDLLVSINHRTL TSGE

CPn_0162 205870 204803
No robust homolog present in Genebank/EMBL as of 11/7/98
FYYTLYNIQSPFRIMKLYSISSDVDTPWIFQLMSKVDSYLFLGGRRIKVVSIVMQEPNLI
IGKVENVRISTIVKILKILSFLIFPLILIALALHYFLHAKYANHLLVSKILERAPGYVPI
MRSGDTASHYKLTTLVPVSOKNLQAMGSNPLEVEAALRTTKPSFFCVPAKYRQIIISSH
LHBCCHENGAGDENIG. TOWNTEVELETMEDE EARDERVENVENCETTTY INCVOKE
LDLKEMGJHLFIDGITGERGALHHTTCOPLITER EVRY LYGHETFGHITT WEGVFFRQG
PLDEDRGGGFEILEQLQELGVRFPICPSGGPDNPNFQGFQGIRIYWEDSYQPNKEV 205870 204803

CPn_0163 205831 206394
No robust homolog present in Genebank/EMBL as of 11/7/98
FEKAIYYCIKCKGIIKCISIIHTPTPATPLCTEGEIFPGPVDSAIQNDLERLLTVKKRPD
IIREYLRAGGSLVTTYPKEQQRLRSPEQLRVLDDLVQSYPNHLHAIELDCGAIPQDLIGA
TYIITFADFSTYILSLRSYQANSPSDDTWGIWFGSIDDPVQAVISFLKDHGFALPSTLAQ

CPn_0164 206444 206998

No robust homolog present in Genebank/EMBL as of 11/7/98

LCFKCIYIKIIFSFLKQLMTRSTIESSDSLCSRSFSQKLSVQTLKNLCESRLMKITSLVI

AFLTLIVGGALIALACGGVLSFPLGLILGSVLVLFSSIYLVSCCKFFTLKEMIMTCSVKS

KINIWFEKQRNKDIEKALENPDLFGENKRNVGNRSARNQLEMILHETDGIILKRYMKGAK

CPn_0165 206983 207592
No robust homolog present in Genebank/EMBL as of 11/7/98
NVLLFNNWVPKTIDHVDPESEIDIRKVVSCYKLIKECQPEFRSLISELLGVIRCGLRLLK
RSKYQEQARTVSDEDAPLFCLTRSYYQDGYLTPLRAGPRDLINHYIHLRRRENPKHFFSP
KHPCYYARLAFNESVCVYRELFDIERLTKMYVEGDYSKEQEKNLQAILSFVKTLDEGKDF LIEHKDTDLIGRGFTDVFCT

CPn_0166 207594 207962
No robust homolog present in Genebank/EMBL as of 11/7/98
NCLKOYNKSDSIMSESINKSIHLEASTPFFIKITMLCESRLVKITSLVISLLALVGAGVT
LVVLFVAGILPLLPVLILEIILITVLVLLFCLVLEPYLIEKPSKIKELPKVDELSVVETD

CER_0167 208309 207977
Nå_Fobust homolog present in Genebank/EMBL as of 11/7/98
NLWSHFPRGFFMLPFCPTILLAKPINSENYGLERLAATVDSYFDLGOSQIVFLSKQDQG
ITWEELSAKDRKFKPGSMNCTLYTEDPILPAKNSFSNCSDIQMRTPISPIH

CPE_0168 208716 208417
Nd :Gobust homolog present in Genebank/EMBL as of 11/7/98
SYNURRRENPEHFFNPGHPCYYARLAFNESVRIYRKLFNTAELKQMYGAGDYEQQNEDN
LKGSLSFVQILDEKDGFDDFLATHKDTTFIGRGGADIFCS

CINED0169 209537 208710

NO. FOUND TO BE THE CONTROL OF THE CONTRO

Defence analog of the transport of transp

CPh 0171 212444 211149 CPNE-0171 212444 211149

*gual-GMP Synthase
IIKLOSARRHLNTIFILDFGSQYTYVLAKQVRKLFVYCEVLPWNSVQCLKERAPLGIIL
SGGPHSVYENKAPHLDPEIYKLGIPILAICYGMQLMARDFGGTTSFGYGEFGYTPIHLYP
CELFKHIVDCESLDTEIRMSHRDHVTTIPEGFNVIASTSQCS:SGIENTKQRLYGLQFHP
EVSDSTPTCNKILETEVQEICSAPTLWNPLYIQODLVSKIODTVIEVFDEVAQSLDVQWL
AQGTIYSDVIESSRSGHASEVIKSHHNVGGLPKNLKLKLVZELRYLFKDEVRILGEALGL
SSYLLDRHPPFGFGCITIRVIGEILPEYLAILRRADLIFIFELRKAKLYDKLSQAFALFLP
IKSVSYKGDCRSYGYTIALRAVESTDFMTGRWAYLPCDVLSSCSSRIINEIPEVSRVVYD
ISDKPPATIEWE

213237 *impD-Inosine 5'-monophosphase dehydrogenase (C00H-terminal region only)
APIGAAIGIGPLGISRAHHLVEAGANVLVIDTAHAHSKGVFQTVLEIKSQFPQISLVVGN LVTAEAAVSLAEIGVDAVKVGIGPGSICTTRIVSEVGYPOITAITNVAKALKNSAVTVIA DGRIRYSGDVVKALAAGADCVMLGSLLAGTDEAJGDIVSIDEKLFKRYRGMGSLGAMKQG SADRYFCTOCQKKLVPGGVEGLVAYKGSVHDVL/YQILGGIRSGMGYVGAETLKDLKTKAS

214041 2137/15 The cobust homolog present in Genebank/EMBL as of 11/7/98
TIFDLIYKIDSYKHOOGFMDFSVFPDRFV#STSPSPIEDIDAKTLVSNCCHYCSRCLFIF
LULLSIIICFSVYGTSGETASLVFGILSYIVLVLLIIECRNRECCRRIS

FVRITESGRAESHIHNIYKVQPTLNY

14724

CPn 0175 2148% 215275 THE TORUST HORIZON PROPORT IN GENEBANK/EMBL AS OF 11/7/98
LILIACE/FILERPEMEQUNCY/ODTTYLYALNSEDERLSDDTHREGKOSPLEAENALGE
FIFSHJYTHSPPLEEVAIPILEOVAFPLYSFIDRDGGGVHYEVLDGVFLKTVAACIIENS FLTDSMSPELLSEVKEALK

CPn_0177 217513 216608

No robust homolog present in Genebank/EMBL as of 11/7/98

DKREQTKSKFIFLISEESMKOPMSLIFSSVCLGLGLGSLSSCNORPSWNYTHNTSTSEEFF
VHGNKSVSQLPHYPSAFFHTOIFSEENDPYVVAKTDEESRKIWREIHKNLKIKGSVIPI
STYGSLMHPKSAAL/TLKTYRPHPIWINGYERSFNIDTCKYLKNGSRRRTSHDCPKNRAVL
NLIKSSGRRCNAILLEMTEEDFVIARREGVYSLYPVEVCSYPQGNPFVIAYAWIADESA
CSKEVLPVKGYYSLWWESVSSSDSLNAFGDSFAEDYLRSTFLANGTSILCVHESYKKVPP
CSKEVLPVKGYYSLWWESVSSSDSLNAFGDSFAEDYLRSTFLANGTSILCVHESYKKVPP

CPn_0178 218052 217789
No robust homolog present in Genebank/EMBL as of 11/7/98
VKEYLDFLYORNVERDPOTKRHCTVSOKFGGESIDAKTTTGQLFHIAGKTEPGHGKLCLG
ESILKQLYALGIITGYENREREVWVYLD

CPn_01/9 218550 218056
No robust homolog present in Genebank/EMBL as of 11/7/98
PKIWDTHFETRIEATSVPKFNRTLRKSFHKSGRSSRPSKACVANFFNFTLQAGRSGIIPG
KKAJLLNVNDAKTPNYSCIFESIGFFNEQDLEAQHNQQAALVRKILKVVPHHFLKGLIAK
LPPSLKKDRKFMSSLIFTKLSYALDLSAPMHLEGKPNLSYEEKLD

CPn_0180 218963 218355
To robust homolog present in Genebank/EMBL as of 11/7/98
TSLHKILDCKYRVFIQNTVASETYPSQILHAQREVRDAYFNQADCHPARANQILEAKKI
CLLDVYHTNHYSVFTECVDNYPNLBFTEVSSKNNEMNGLSNPLDNVLVEANVRRTHARNL
LAACKIRNIEVPRVVGLDLRSGILISKLELKQPOFOSLTEDFVNHSTNQEEARVHQKHVL LISLILLCKQAVLESFQEKKRSS

CPn_0181 219175 218777
No robust homolog present in Genebank/EMBL as of 11/7/98
VHELFKIDGVYYFFKKFMKLFYNNYSLNSHHEKPSSLEKAVQALDSYFYWGGDTTDVLAR
DDISREIYCVRRLYINFWIVSISQSLSRIPWRLKRILLRYCTLRGKYVMPILIKRIAILL
GITBFSBLEYGVV

CPn_0182 220704 219334
accC-Biotin Carboxylase
RCIMKKVLIANRGEIAVRIIRACHDLGLSTVAVYSLADOEALHVLLADEAICIGEPQAAK
SYLKISNILAACEITGADOAVHGYGFISENANFASICESCGLTFIGPSSESIAMMGDKIA
AKSLAKKIKCPVIPOSEGIIEDESGLKIAEKIGFPIVIKAVAGGGGRGIRIVKEKDEFY
RAFSAARAEAEAGFNNPNYYIEKFIENPRHLEIQVIGDTHGNYVHLGEBCCTIORROKL
IEETPSPILNAEIRVKVGKVAVDLARSAGYFSVGTVEFLLDKDKKFYFMEMTHRIQVEHT
ITEEVTGIDLVKEGIHVAMGKLHPMCOKNIEFSCHIIQCRINAEDPTNNFSPSPGRLDYY
LPPAGPSIRVDGACYSGYAIPPYYDSMIAKVIAKGKNREEAIAIMKRALKEFHIGGVQST
IPFHQFMLDNPKFLESNYDINYIDNLLAQGNSFFKEF

221207 220695 CER_U103
accB-Biotin Carboxyl Carrier Protein
RRICMDLKQIEKLMIAMGRNCMKRFAIKREGLELELERDTREGNRQEPVFYDSRLFSGFS
QERPIPTDPKKDTIKETTTENSETSTTTSSGDFISSPLVGTFYGSPAPDSPSFVKPGDIV SEDTIVCIVEAMKVMNEVKAGMSGRVLEVLITNGDPVQFGSKLFRIAKDAS

CPn_0184 221814 221221
efp-Elongation Factor P
OWKIKFCCCEEKIMVLSSQLSVGMFISTKDGLYKVTSVSKVAGPKGESFIKVALQAADSD
VYLERNFKATQEVKEAGFETRTLEFLYLYLEDESYLFLDLGNYEKLFIPQEIMKDNFLFLKA
GVTVSAMVYDNVVFSVELPHFLELMVSKTDFPGDSLSLSGGVKKALLETGIEVMVPPFVE IGDVIKIDTRTCEYIORV

222457 221765 rpe/araD-Ribulose-P Epimerase AEVKKQESVLVGPSIMGADLTCLGVEAKKLEQAGSDFIHIDIMDGHFVPNLTFGPGIIAA INRSTOLFLEVHAMIYNPFEFIESFVRSGADRIIVHFEASEDIKELLSYIKKCGVQAGLA FSPDTSIEFLPSFLPFCDVVVLMSVYPGFTGQSFLPNTIEKIAFARHAIKTLGLKDSCLI EVDGGIDQQSAPLCRDAGADILVTASYLFEADSLAMEDKILLLRGENYGVK

222878 224068 *similarity to Cps IncA
PIKDKILMSSPVNNTPSAPNIPIPAPTTPGIPTTKPRSSFIEKVIIVAKYILFAIAATSG
ALGTILGLSGALTPGIGIALLVIFFVSMVLLGLILKDSISGGEERRLREEVSRFTSENQR ALGTITGLEGGALTWIGIALUTFFVSMVLLELLKUSTSGGEERKREEVSKTSENVK LTVITTTLETETVRLKAARDQLTLEIFEAFRNENGILKTTAEDLEEQVSKLSEGUEALERI NQLIQANAGDAQEISSELKKLISGWDSKVVEQINTSIQALKVLLQQEWVQEAQTHVKAMQ EQIQALQAEILGMHNQSTALQKSVENLLVQDQALTRVVGELLESENKLSQACSALRQEIE KLAQHETSLQQRIDANLAQEQNLAEQVTALEKMKQEAQKAESEFIACVRDRTFGRRETPP PTTPVVEGDESQEEDEGGTPPVSQPSSPVDRATGDGQ

CPn_0187 224218 225045
predicted methylase
vPLTYTRTLPMHSKFLSRRKKNSSHKEETCWDCIASSYNKIVODKGHYYHRETILPOLLP
SLTLOSKSSVLDICCGOGFLEPALPKECRYLGIDISSPLIALAKKMRGVNSHOFKVADLS
KRLEFVEPTLFSHAVAILSLONMEFPSFAIRNTATLLEPLGOFFIVLNHRCFRIPRASSW
HYDENKKAISRHIDRYLSPHKIPIMHIMOKOBOPTTLOFHFPLCYWFKELSSHOFLVSGL
EEWTCSKTCTGKRAKAENLCRKEFPLFIMICCIKIK
CTn_0188 25090 226400 224218 225045

CTI 12 hypothetical protein CT EMPERATE AND ADMITTANCE OF THE PROPERTY OF THE ADMITTANCE OF THE PROPERTY OF T

FIFLODKILPSCYLQLITSTYILALTTROFNEGLSPLT EKECFLFFYNNCYOPVFNFSELTIKDIADKLLHREIFK SKNKENLTLGEIARRIK

OSKVPIGEVSQCLDVL GITFIENSFQNIFNQA

CPn_0189 226391 229825
CT131 homolog (Possible Transmembrane Protein)
ANOMKRESMIKIIGICIASSIVIGFLIFLPQLLSTESRKYLVFSLIHKESGLSCSAEELK
ISWPGROTARK IKLTGEAKDEVFSAEKFELDGSLLRLLIYKKPKGITLGGWSLKINEPAS
HARDWSHLDHOLDLE SERVELDE VAR FOR STELLEY FROM THE HER FLOWER FOR THE TRANSPORT OF THE PROTECT OF T

CPn_0190 229901 231274

No robust homolog present in Genebank/EMBL as of 11/7/98

LLGIKLMRKRHSFDSTSTKKEAVSKAIQKIIKIMETTOPSLAVETPNAEIESILQEIKEI
KQKLSKQAEDLGLLEKYCSQETLSNLENTNASLKLSIGSVIEELASLKQLVEESIEESLG
QDQLIGSVLIEISDKFLSSIGETLSGNLDMYQNVIQGLLIKENPEKSEAASVGYVQTLL
EPLSKRIGETHKKVATHDVNISSLQFHMMSVAGGRFRGHIDMYGYRVLGLGEPKNGEDAV
SKDYLERYVSSQLTIDKVEDKPITKPNKGKLLYSQGTSPKLESPLPLGLLTSGISGFTWK
SASKSNGSSFPFSALRIKETESDTDCFQITSTTLGGNQAGTYTMSLSLKVLVPSIFQIEK
PEVQLSLVYSYEDWLPIDNIFNMSQPRTIPLALLCQTMLAGQKYDILELAAHQTNQTLMI
SPNCSRFSLQLKQTNQFENSPVDFYIVHAAHSCHWSGF

CPn_0191 232039 231314
ging-abc Amino Acid Transporter ATPase
LEALAGLVQPTQGDIWIEGEAPALIVFQOPELFSHMTVLGNCTHPOIHTKGRSTEEAREKA
FELLHLDIEEVAKNYPDQLSGQKQRVAIVRSLCMDKHTLLFDEPTSALDPFATASFRH
LLETLRDQELTVGLTTHDMQFVHSCLDRIYLIDQGTVAGVYDKRDGELDSGHPLSKYIHS
AQ

232643 231984
ghp-abc amino acid transporter Permease
bygydhlafarillegggyttevsgigilgsilglitgtvtslyfpskltkllansyv
tvtrgtplfigiliiyfglpsvlpieftplyagiialsknsaaylaenirgginslsig
wesamvlgykkygifyiiypgvfknilpsltnefvslikessilmvvgvpeltkvtkdi
vsklnymemylicaglyflmttsfscisrlsekrrsydn

CPR 0193 233144 232686
*argR-Arginine Repressor
KLHGVFMKKKVTIDEALKEILRLEGAATQEELCAKLLAQGFATTQSSVSRWLRKIQAVKV
AGERGARYSLPSSTEKTTTRHLVLSIRHNASLIVIRTVPGSASWIAALLDQGLKDEIUST
LAGDDTIFVTPIDEGRLPLLMVSIANLLQVFLD

CRILO194 233162 234241
gGF-0-Sialoglycoprotein Endopeptidase
EVERTIKGNYFFSWFFMLTLGLESSCDETACAIVNEDKQILANIIASQDIHASYGGVVPE
LASKAHLHIFPGVINKALQQANLLIEDMDLIAVTDYPGLIGSLSVGVHFGKGIAIGAKKS,
LLGTMHVEAHLYAAYMAAQNVOFPALGLVVSGAHTAAFFIENPTSYKLIGKTADDAIGET
FDRYGRFLGLPYPAGPLIEKLALEGSEDSYPFSPAKVPNYDFSFSGLKTAVLYAIKGNNS
SHESPAPETSLEKQRDIAASFOKAACTTIAOKLPTIIKEFSCRSILIGGGYAINEYFRSA
IQTÄCNLPVYFPPAKLCSDNAAMIAGLGGENFQKNSSIPEIRICARYQWJSVSPFSLASP

CPH_0195 234172 235785

oppA-Oligopeptide Binding Protein
YSGNSYMRKISVGICITILLSLSVVLOGCKESSHSSTSRGELAINIADEPRSLDPROVRL
LSEISLVKHIYEGLVQENNLSGNIEPALAEDVSLSSDGLTYTFKLKSAFWSNGDPLTAED
FIESWKQDVATQEVSG IYAFALNPIKNVRKIGGGHLSIDHFGYHSFLSVVLTLESPTSH
FLKLLALPVFFPVHKSQRTLOSKSLPIASGAFYPKNIKQKOWIKLSKNPHYYNOSOVETK
TITHFIPDANTAAKLFNQGKLNWQGPPWGERIPQETLSNLOSKGHLHSFDVAGTSWLTF
NINKFFLNNMKLREALASALDKEALVSTIFLGRAKTADHLLENNIHSYPENGKQEMAQRQ
AYAKKLFKEALEELQITAKDLEHLNLIFFVSSSASSLLVOLFREOWKESLGFAIPIVGKE
FALLQADLSSONFSLATGGWFADFADPMAFTLIFAYPSGVVPYAINHKDFLEILONIEGE
ODHOKRSELVSQASLYLETFHIIEPIYHDAFQFAMNKKLSNLGVSPTGVVDFRYAKEN

CPn_0196 235906 237519

oppA-01igopeptide Binding Protein
KLKSYSKERSFMLRFFAVFISTLWLITSGCSPSOSSKGIFVVNMKEMPRSLDPGKTRLIA
DOTLMRHLYEGLVEENSONGELKPALLESYTISEDOTRYTFKIKNILMSNGDPLTAQDFV
SSWKEILKEDASSVYLYAFLPIKNARAIFDDTESPENLGVRALDKRHLEIQLETPCAHFL
HELTLPIFFFVHETLRNYSTSFEEMPITCGAFRP/SLEKGLRLHLEKNPMYHNKSRVKLH
KILVOFISNANTAAILFKHKKLDWOOPPWOEPIFPEISASLHODDQLFSLDGASTTWLLF
NIOKKFWNNAKLRKALSLAIDKDMLTKVVYOG/AEPTDHILHPRLYPGTYPERKRONERI
LEAQOLFEEALDELOMTREDLEKETLTFSTFSFSYGRICOMLREOWKKVLKFTIPIVOQE
FFTIOKNFLEUNYSLTVNOWTAAFIDFMSYLMIFANPGGISPYHLODSHFOTLLIKITQE
HKKHLENDLIIFALDVLEHCHILEPLNHRAFIAFNSFRFIEKL

CPI_0197 237512 2/8882
opps of igopopticle triding Post-tin
knyrrk; iddkfdokfetky femerby/theflet iddn/cooysokhkobili (Pihddey
kfberdakramulo: Laglefo lotrophresholdela Labryty-edfecoytefikobal
wodotfettegolkniaweyaqensph/q (pogline): ppsnattildoshphdepkllafe;
kfa fekrenykleodytytuveyepini hilkknypryydyhoy insiklli (polytyta)
lankovuwwydharo; it welhiyo; oynyytyty pyggafwilolnyksphlindlonrirla
lankovuwwydharo; it welhiyo; oynyytyty pyggafwilolnyksphlindlonrirla
etilke, wkaacitoli liegileyyl-ynkrkydyya i attythyayytgaali i seedklunk

ELIPTYYLGYDYLTODFIE

CPN_0198 237 24076

OppA-Qligopeptide Binding Proffin

QIEYYIMKMHRLKPTLKSLIPHLLFLLUFLSSGSKOKQEPLGKHLVIAMSHDLADLDPRN

QIEYYIMKMHRLKPTLKSLIPHLLFLLUFUSGGSKOKQEPLGKHLVIAMSHDLADLDPRN

AYLSRDASLAKALYEGLTRETDQGIALALAESYTLSKDHKVYTFKLRPSVWSDGTPLTAY

DFEKSIKQLYFEBFSPSIHTLLGVIKMSSAIHNAOKSLETLGIQAKDDLTLVITLEOPFP

YFLTLIARPYFSPVHHTLRESYKKOFPPSTYISNOFPVLKKHEHONYLLLEKNPHYYDHE

AND OPPTLLVIIPHAGATATVLFKOV INMICIPHOAPISNEFOKVISGERILTVSVSTT

BLITTHLAPPLIOHFALAF LABBAY & ELIBERTSFORMATOLFFHLOGHLFEIGTEZ

ROTKARAVYFDEAKETLEFELABLISTET IDJONDITIAGETJAGGKETTGJETTGGME

YHCFLKKROGDFFIATOSWIAETVSPVAFLSILGNPROLTOWRNSDYEKTLEKLYLPHA

YKENLKRAEMIIEEETPIIPLYHCKYIYAIHPKIQNTFGSLLGHTDLKNIDILS

CPn_0199 2/1018 241983
oppB-01igopeptide Bermease
KCLIGLSLVFSYIKNRILFNLLSLMIVLTLTFLVMKTIPGDPFNDEGCNVLSEEVLQTLK
SRYGLDKPLYQQYTQYLMSIAKLDFGNSLVYKDRKVTNIISTAFPISAILGLQSLFLSIG
GGIALGTIAALKKKORYILGASILQISIPAFIFATLQYVFAVXIPLLPIACWGSFTH
TILPTLALAVTPMAFJIOLTYSSVSSAINKDYVLLAYAKGLSPLKVVIKHLIPYAIFPTI
SYSAFLTTTVITGTSAIENIFCIPGLGKWFICSIKQRDYPVALGLSVFYGTLFMLSSLLS
DLIQSIIDPQIRYAIGKEKKRK

CPn_0200 241996 242868
oppC-0ligopeptide Permease
EKKKKKIMENISSAPSRSIWKSIIONKMLVLGLTTLIILMLGALLLPWFYQDYECTSLKD
ILVSPCSRFPFGTDTLGRCMFARTLRGLRISLLIATIATLIDVCVGLLWATVAISGGKKI
DFLWRITEILFSLPRIPIIIILLLVIFHHGLLPLILAMTITGHIPISRIYGGFLLKNI
PFVLSAKAMHASTFHILKKHLLPNTLAPIISTLIFTIPNAIYTEAFISFLGLGIQPPQAS
LGTLVKBEINAIDYYPWLFFFPSLIMIALSISFNLIGEGAKTLCLEEGSHG

CPn_0301 242810 243715

oppD_01igopeptide Transport ATPase
asis_ARGLKHYVSKRDLMGNYLINIKDLTITSTNPKRTLIENLSLQLKENRNLALVGES
ASIS_ARGLKHYVSKRDLMGNYLINIKDLTITSTNPKRTLIENLSLQLKENRNLALVGES
GSGKTTITKAILGFLPENCLIKTGSILFEDIDITKLSPKELHKIRGGKIATILONAMGSL
TPSMRIGMGIIETLROHHKMKERAYNKAMQLLTDVCIPNPKYSFSQYPFELSGGMRQRV
VJATALASOPKLILADEPTHALDSNSQAQVLRILRNIQOGKQATILLVTHNLSLVKELCN
PICIIKDGKLIETGTVEEIFLSPKHPYTLKLLNAVSKIPIKKTSSPILKNKFQPLMSMOG

CPn_0202 243682 244500
oppF-0ligopeptide Transport ATPase
oppF-oligopeptide Transport ATPase
vPTSNEYARMFMTTLLSIKDLSLTIRGKKILNHINLNLIKGSYLTIVGPSGSGKSSLALT
ILDLLKPTTGTITFHMDPKIPRARKVQVIWODIDSSLNPCMSIKGIISEPLNIIGTYSKA
EQNKEIYNVLDLVNLPKSVLHLKPYKLSGGQKQRIAIAKALVSKPELLICDEPLSSLDTL
NQSLILDLFQTIKKEYQNTLLFITHDMSAAYYIADTIAVMDQGSLVEHACREKIFSTPKH
TTTQDLLDAIPIFFSLISTEMEPSEEYELQVASK

CPn_0203 244966 245802
No robust homolog present in Genebank/EMBL as of 11/7/98
IVPLPQKNNKETSCMNTYTFSPTLQKSFSLFLLEKLDSYFFFGGTRTQILVITPTNIRLA
AKKRGCKVSTIEKIIKILSFILLPLVIIAFILRYFLHKKFDKQFLCIPKVISNEDEALLG
SREQAVEKAVREISPAFFSIPKKYQLIRIDTEKDDAPSILFPIGIEIILKDLCIDTLKQS
NLFLKREMDFLGHPEEKALFDSICSIEKDQEWMSLESKKLLITHFLKYLFVSGIEQLNPG
FNPENGRGYFSEISTAKIHFHOHGRYGPIRSSGPIMKEI

CPn_0204 245691 246002
No robust homolog present in Genebank/EMBL as of 11/7/98
PREWAMVFFRNKYSKDPFSSARSIWANPFFGTHHEGNIKIKGMGYQIFTRLKKLGISFSS
YNSINPMPYFFDECCFVYWESOFKSALDDHGILQKQTETFYRNT

CPn_0205 246073 246327
No robust homolog present in Genebank/EMBL as of 11/7/98
IEDSIKGYGSASAFRNPPQLLLKFFLVCEELCILTVATHRALLETPLALSFFKELKTKYV
YRAKDILQLHNYKGFTILNTSPLCS

CPn_0206 246346 247161
CT203 hypothetical protein
IVDRRSPACYDSINSDAIGVSLLMDISHILEDLAYDEGILPREAIEAAIVKQMQITPYLL
HILHDATQRVPEIVNDGSYCOHLYAMYLLAQFRESRALPLIIKLFAFEDDTPHAIAGDVL
TEDLPRILASVCNDDSLIKELIETPKINPYVKAAAISGLVTLVGAGKIPRDKVIRYFAEL
LNYRLEKQPSFAWDNLIAGICTLYPGELFYPISKAFDCGLVDTSFISMEDVENIIHEETV
ESCIHTLCSSTELINDTLEEMEKWLEDFPIEP

CPn_0207 247209 248617
ybh1/sodit1-0xoglutarate/Malate Translocator
vnkkkrflsllfltavlljiwfsphpasinsnawQlfaiftttimgiifopvpmgaiaii
gistllitottileQclsgfinpiawLvflsfsiakgiiktglgeriayffvsalgkspl
glsyglvttopfflapaifsvtapacgilypvvtslsdsfgssaekctqoligsflikvay
ossvitsamfltamagnflvaalaghvgvslswvlmakaaiipolpslflmpiilyklyp
pkitoceeairsakLrlexpplkkeektilmiffluv/Lwffgbllgisattaaligls
LliltnildbwckDvtanttametfftwfgalimmasflnqlgfiplvcdsaaalvsglswk
igfpllfliyfyshylfasntahigam/piflavsislctnpifaaltlafasnlfgglt
hygsgpaplyfgshlvtv?emwrsgfalsivniviwigigslwwkalgli

CPI_0208 24935 250602
pfkA-Fructose-6-P Fhosphot ransferase
svaviLmMPLVVDLDTD IISVSPPLPKEPQEAASLIAVPDTSHSKPVVPQVKTLFPQTYH
LPYLKFVQGENVVHTPLKVGVMFSCGPAPXGHNVIOGLFNSLKDFHPDSSLVGFVNN3DG
LTHING IDITEEFLSKFSNSVGFNC IGTGYRK I VTPEAKEACLKTAEALDLDGLV I IGGD
GOTHTATA I LAEVFAKRENTSH I VGVPKT I DOLOPTEPLDITEGFDTATKFYSS I ISN ISR
DALOCKAHYHFIKLMSRASHIALEGALOTHIBITAL I GELAEKNDELKTI I HKICSVIA
DRAMEKYYGVI I I FEATI SEFIPET I HILTET I EGISTEREN PEDRI DECQLLKSFPAPI
LEGITINDRAHGNOVYN KOVDVLULHILVSNIHLDGVYFPIVPFNA I SHELGYEGRSGLPTK
FORTTO ZYSLOYGAGILANNIE SEVYLGTI I GELAG PPRIVMLRA I GVVKMFTVKQQADGTLQ
PK I PKYLVD I GGTAFRKEKEYPY I WALEDSYPPIJ JEDGYTET PEMHEDNFPPLTELLINHN
FWORIOG SCIET POTTY

CTm_0209 251999 251272
No cobust homolog present in Genebank/EMBL as of 11/7/98
NSCHMIRKMETYSFSTEL&KHTSLYIMEKLSCYFSFG/SKRTRVIAUTPAGLAIAYEONI

HESMTYK LEKYESFPRSEERTTSEWYRP

CPn_0210 252459 251440
No robust homolog present in Genebank/EMBL as of 11/7/98
YQKLWEREREYFKTIREKEHATISTMLVELEALKREFAHLKDQKPTSDQEITSLYQCLDH
LEFVLLGLGGDKFLKATEDEDVLFESQKAIDAWNALLTKARDVLGLGDIGAIYQTIEFLG
AYLSKYNRRAFFILASEIHFLKTAIRDLNAYYLLDFRWPLCKIEEFVDWGNDCVEIAKRKL
CTFEKETKELNESLLREEHAMEKCSIODLORKLGDIIIELDDVSLFGFKTPSQEEYQKD
TAGGERERELEGETTEL FERTETDE FOR AANARFIREFESTALEGETTE STOPPOTKELEFAI
ARJELEGERELTWEEYERAAAEGULTSWEEESTALEGETTE SENTRE

CPn_0211 252765 252463
No robust homolog present in Genebank/EMBL as of 11/7/98
ECVMSYPDISNVQASSIQSALLHKTSDQIQOKRCFKQSTFVILAVSLVIIGSLFLLAGVA
ILTVFSHGVLSLVFGVLGIVLGLLLLAGGVGLLVEEAKSLI

CPn_0212 254081 252888

No robust homolog present in Genebank/EMBL as of 11/7/98
ELSYGVVVSIYSEILSFSELTSCKHSLFPFGPIETASIRIHHVFNVVIVCLIILGTLFVC
LGMVFLGVFSTYLLGMSSMILGILLISIGLALLKFKERYGLEPKELFGVEGGFDKKLPSE
IIOMODQIADLARELDLEQKKDTLIRGFSARLDVLEGSKTEKKQILKIGVPRNLSEIQER
AQEONSILEÇCKEALLFRRKSAQEIFKKLYDRKAAFWRSYREDLMCYSEIHVSKKALSNL
YIGDVFEGTAPHFLMEAYAMCRTAKNLRNYVKVCVEDMRVNEEKKRAKQLSVSELLCCCT
EIETDLENETNLFTSDSEDVLEEYQIHCIRVTMLHALWAIYNDEVVSRKPIDTLDRVRAR
MAVEDCIETFEELQMCVVHTKTLELEIAQLYVDILLEA

CPn_0213 254345 254190
No robust homolog present in Genebank/EMBL as of 11/7/98
ILVVFSRVIFSNTNQIGIPRLELILPLWKKENDPFCFLFSRVEGTFIILNIK

CPn_0214 255768 254446
No robust homolog present in Genebank/EMBL as of 11/7/98
FLGLKEDYERPTYCHI PPAPHPORVDSKGCIASHVSTVVVVALFILGIFFLSGSLAFLVH
TSCGVLLGAALPILCIGLVLLAVALIVFLCHKHKTRQDLDYYDQDLDSLVIHKKEIPNDI
SELRVIFEKLQNLFOFHTKDFSDLSQELQGKFINCMERKLILLEDEVITKFLIVDRFLETR
RNFTTFGEOVKGIQSHIFDLHEEKSLVLELYRLRRDLQVLLMFFLLPPGILKVDYDEIE
AIKGLFIRLTSRLDKLDVKAQERKKFINEMSREFKEVEKAFDIVDRATKKLMDRAKKESP
ARLFWGRTESLLEMKNEEALKNQGLDFENLSHPELFSPYQQLLILNYLNSEIVLHHYEF
LISGTVTSGLUTLEECENRWRAASTGLNALLVRKLQFRGAIKSAYFEKLTEIEKELRSLQD
VINSLELELIHKIKDIVTEET

CPT_0215 257039 255759
No.Tobust homolog present in Genebank/EMBL as of 11/7/98
LT\$\$KKQVMSSATARDCFFSPSPQPSSTLGVMPPKYKSLILSVSLIVLGVLLLCVGMLLL
VMAFFFSFSVLTVGLGGAGVFLGSLLLILGLIFFVSYHRKLSEATRSLEQKTILEYQPWAD
LR\$ELNEVQEWSNFLLDEWEDFKEVVAQHKSQFATFEGDLLLFGREVEKYETIMKELDGR
DVABLTELKNIWGPLEFLRKKGDRLQCEIDKLRKEVMKVGKSGLKLACELTKFKSALKDV
KLEGGECYROKRKVEKLEVPFEGYRRELLEVLKTRLSVEKEQLFEEVVSAFERLASLHR
TVF\$SEELQBALDKAKAELLDIQVRKSVVEDLSCEPTLIQVHLLRLYEVQCRIVEQFLTQ
TFSSEQEKVLEEYEALKARIRKTLRVKLDQVRANVAFVASTTDLLSESESLDGNDSVFED
AMDDFLD

CPT 2026 257623 257174
NO FIDURE HOMOLOG PRESENT IN GENEBANK/EMBL as of 11/7/98
NKÄRTMNPVTFDRIQVDFIPEDTSLRINSYIVAGGLILGVVLSILSVICLDIGLVGLSA
GAAFTLGLGCLIFALFLFSFSLILLLSQEKRVPDVLSLYLEKEVPQYETPLYKEDLESER
DIŞAISERLGIIEEKLRIAEKFRYSDSVFV

CPn_0217 257881 258579 ypdP

ypdb P PKECKLKGFLSVNELIFGFQTFSVVVLGVFFASRGKAWLTGWLSLLSSIMNVFVLKQFHL WGFEVTSADVYVIGLLTCLMYAREHYEKNDINDAMLCSWISIAFLVLTQLHLFLIFSPN DS\$QEHFLALFSSTPRIVVASLVTLIFVQIVDIKLFTFLQRVFSKKYFAMRSTISILFSQ LIDIIIFSFLGLYGLVSNLCDVMIFAMLVKGIVITLAIPTLTVTKAVLDRRSS

CPH_D218 259064 258582
NOLEGOUST HOMOLOGY PRESENT OF THE PROPERTY OF THE PROPE

CPn_0219 259348 260472

tgt-Queuine tRNA Ribosyl Transferase
GSSLALKFHLIHOSKKSQARVGQIETSHGVIDTPAFVPVATHGALKGVIDHSDIPLLFCN
TYHLLLHPGPEAVAKLGCLHOFMGROAPIITDSGGFQIESLAYGSVÆEIKSGGKKKGMS
SLVKITDEGAMFKSYROGRKLFLSPELSVQAOKDLGADIIIPLDEL PFHTDOEYFLTSC
SRTYVMEKRSLEYHRKDPRHQSMYGVIHCGLDPEQPRIGVRFVEDLFPFDCSAIGSLGRN
LQEMSEVVKITTSFLSKERPVHLLGIGDLPSIYAMVGFGIDSFDGSYFYFKARAHGLILSK
AGPIKIGOOKYSODSSTIDPSCSCLTCLSGISRAYLPHLFKVREPNAAIWASIHNLHHMQ
QVMKEIREAILKDEI

CPn_0220 260660 261238
No robust homolog present in Genebank/EMBL as of 11/7/98
FYSFLKKKGIFYMSKESIRSYSEISTPTPIFRETPSKEGVAYKLOLRSPAKDCILRNRVS
LKGALLRSIFFYGSFLGAKRIHSAMSAKDAPCTTRV/HYLVGGLELLGLGVVVLACKVLA
TALKFLFSKASSKIKQMKWREKARNLAAKDTVQSIKEFOSVDLTSCFTRCFRLRNRVVEE
GASENOTVREIIV

CPh_0221 261621 262064
NO tobust homolog present in Geneburk/EMBL us of 11/7/98
DALRYKYELGIQMVNRYKSASEFSADHYDDHLVBYGYKRNLRGLAPVENEVCLFEENNL
LESYMAGIFIMGHILGUGRLHSVWGTODPKDSKI/HIFHTALGILETUNGHIVLLIKHT
LTHLLILFTPOLLUYFMYGAAYSDFHPI

CPH_00229 262474 20284/ -Weak dimitatity to Bacteriophage CHP1 (Ott4) GEKELKIWEKLRELNAFELTOPEEYRHRWYMMTGLEGREGREGHAKWESYRCVHEASLYE ENDFETTITYDDKHLPQYGGEVKLHLOLFLKELRKMIGPHKIRYFEXDAYGTKLORPHYHL LLC

CPn_0223 262950 √63333

No robust homolog programmer in Genebank/EMBL as of 11/7/98 intmLigryscoopertants, ktdpvnonLegutneisetvsetsssikbsvlnsleidsiloganlystustnopedetgekiwhtifgaletleleilillekiifvllheifHLVigfeK

CPn_0224 263402 263674
No robust homolog present in Genebank/EMBL as of 11/7/98
YTEKNPKKNKKMKPNSIFFLENTKHYPDIFREGFVRDRHGLMEASDWLLSTEITIIRSIL

CPN_0225 263946 264541
No robust homolog present in Genebank/EMBL as of 11/7/98
NSFTIKFLLMTKNAINSOTTF POPNLTDAEPIASRACKSIAVIISLFALGMLLLCLGII
LISIPIPGLAAQVALGLGIYSLILGIALANIGFLCLLLRCKOYPOPDTLPSESSKOPSE
GSTETALDWOAGEFLEKVOVSATPILLPKNKDEELSAKVMKEGAEAASSIKQAVLESTEK
LIDARKQEESRREARKKIVAEEAEASRKRIQOQMAADQEALRKRKEEVAKRK

CPn_0226 / 264545 264967
No robust homolog present in Genebank/EMBL as of 11/7/98
AIFNRKRMPYYANTLEFIQGTQSLCPLFKYGFVRHYKGQLEIEDASHDWDFLEPPSTWK
RTLLAAIPILGSVIĞLGRLFSIWSIREPQDSQEYKSIFWHTLCAVLEILGLGIVALILKI
LATFIMAMPGLKRYATFLFYS

CPn_0227 265467 265009
dsbB-Disulfide bond Oxidoreductase
KERFNIFVSCHLLKEIMMINFIRSYALYFAWAISCAGTLISIFYSYILNVEPCILCYYOR
ICLFPLTVILGISAYREDSSIKLYILPOAVLGLGISIYQVFLQEIPGMQLDICGRVSCST
KIFLFSYVTPMASVVAFGAIVCLLVLTKKYRG

CPn_0229 266242 265412
dsbG-Disulfide Bond Chaperone
VKDRADFINLKEKFSCSILKKENAFEFYYFCSIKQLTNSSLRGPLNKKILVLCTAMFFIV
CFGFLJHKKHTILPPKAHIPTNAKHFPTIGNPYAPINITVFEEPSCSACAEFTTEVFPLL
KKHYIDTGEISFTLIPVCFIRGSKPAAQALLCIYHHDPRQADIDAYMEYFHRILTYPKEE
GSHWATPEVLTKLAEGLKINSGRSVNPKGLEQCIASGQYNEQIKKNNLYGSQVLGGQLAT
PTAYVGDYLIEDPTFHEIERAIQHIRQLQAVEGDHDD

CPM_0229 266163 267560

CT178 hypothetical protein

#SKAFSFLRIEQENFSFKFKKSALSFTYNTANLTKSTFTFILLLLERKDOGLRFMDKET

LENIYRHFRYRFLKLNILPAFLGLLLLCSPNTLNYTQVDVIFSDRLCSCLLIFLAIASLT

KRSLLWLGAPLGIWYTLFACVAGRSPTIFANDTLIGFAILAVVCISFTRPEALEVGPTLP

EGFSYNPSAGGRRAAVLFLSLLGWLEARYLTASSLGITSSQSSNFLLLYSSIMTVYSLLV

VLSLAGSERRWHTRPKIVIATALALTGVIILTLLPIILLQLRYDCWLCLCLTIEPALAVV

FAYDETRATLRYISQFLGDKRALTRASFFGSEYYKHTLSWEERTVLPLRKAYKQAFBGIS

FPINQLLAILVATVFVKVNSSMGLPTFPRNFLNICCWFIIVLFILAFAESLRHLRWMNLI

SAAAILFSPVLFHIPVESPMFLPIIVTGLILIILSIGKRRTKKKL

CPn_0230 268277 267576
CT179 hypothetical protein
RFKKALIYMSSOPLVTTSSSLSRYVVLTGEEKVACYKKAFNHIWHGAPAIILAAALLMFC
IFGFVLGSILLGAPLEGASILYDVILPWLLPSILVFVLLVLPLNIYAYSHHKQVLALHER
ITOSNYKEIYDHCEKEKKTPNKKALSLYIESQVLVPEYSKRFSSMILGKTLKIIPKKDSP
ESLKHDELIQKALERAKENIYMNKNQREKRDEREAKKEAKNASKTNPLWEGLGT

CPn_0231 268996 268253
tauB-ABC Transport ATPase (Nitrate/Fe)
PQARVSIQDAGFSHLQAHRLCYSCONQVILKDASFQASPGTITIILGSSGVGKTTLFRLL
AGFLPLQBGELLWNSSPLNRKDVAYMQQKEALLPWRTALKNMTLSTELGINTSHNALSNE
RLEEIIHNFDLGQLLDRYPDELSGGQRQRIALAAQCLSLKPILLLDEPFSSLDVLLKEQL
YQDIVALAKKENKTVLLVTHDFHDVSCLGDVLYVIKNKTLTPVPLDPSMRPLANGLCFIK
DLKKHLTY

CPn_0232 270134 269232
*similarity to 5'-Methylthioadenosine/S-Adenosylhomocysteine
Nucleosidase
KKFLMRRFLFLILSSLPLVAFSADNFTILEEKQSPLSRVSIIFALPGVTPVSFDGNCPIP
WFSHSKKTLEQQRIYYSCDSFGKYFVVSALWPNKVSSAVVACNMILKHRVDLILIIGSCY
SRSQDSRFGSVLVSKGYINYDADVRPFFERFEIPDIKKSVFATSEVHREAILRGGEEFIS
THKQEIEELLKTHGYLKSTTKTEHTLMEGLVATGESFANSRNYFLSLQKLYPEIHGFDSV
SGAVSQVCYEYSIPCLGVNILLPHPLESRSNEDWKHLQSEASKIYMDTLLKSVLKELCSS

CPn_0233 270439 270248
No robust homolog present in Genebank/EMBL as of 11/7/98
EKARTMFLGKVLLFLLRISRRSYVQEIGIFFHLETPDLKIVLCAFVSTFIVVEMDVSLKN
KCOS

CPn_0234 271246 270548
CT181 hypothetical protein
FIMLOSCKKALLSIVVSILAFHPIPGMGVEAKSGFLGKVKGWFSKKEIQEEARILPVKDS
LSWKRYDYTSSSGFSVEFPGEPDHSGQIVEVPOSEITIRYDTYVTETHPDNTVYVVSVWE
YPEKVDISRPELNLQEGFSGMMQALPESOVLFMQARQIOGHKALEFWIVCEDVYFRGMLI
SVNHTLYQVFMVYKNKNPQALDKEYEAFSQSFKITKIREPRTIPSSVKKKVSL

CPn_0235 271395 272177
kds8-deoxyoctulonosic Acid Dynthetase
VFVRYLLMKPEESECLCIGVLPARWNCSRYPCKPLAKIHGKSLIQRTYENASOSSLLDKI
VVATDODHIIDHVTDFGGYAVMTSPTCSINGTERTGEVARKYPFKAEIIVNIQGDEPCLNS
EVVDALVQKLRSSPEAELVTFVALTTDEEEILTEKKVKCVFDSEGRALYFSRSPIPFILK
KATPVYLHIGVYAFKREALFRYLOHGSTPLSDAEDLEOLRFLEHGGKIHVCIVDAKSPSV
DYPEDIAKVEQYITGLSNAYF

CFIL_0236 272138 273766
pyco-ctp Synchediag
birty-impekcifeld/groupolgkglit/aglalilerorlinvamekedpylnvdfotmnp
fergelyvtddrivetdedolghyhregozaalghiogatsog fyarvikreregdylgstvo
vihittre fovildaakehopvelt/effyttigd eglefelea irofrydhoedenih
mtyvpyldaadevkskproheyvdth/gryfiipdiederprofeykreisleronypn
avifnytdvametyempimeagekiahfioekkreindevkreisleronypn
avifnytdvametyempimeagekiahfioekkreindevkreisleronypn
avifnytdvametyempimeagekiahfioekkreindevkreisleronypn
avifnytdvametyempimeagekiahfioekkreindevkreisleropocket
gvyckyvohrdayksiffealthaalenihaafiipidaedenlitmelgocdaclypgofg

VRGWEGK I AVAKFOREGG I PYFG I CLGMGVLVVEYARN VMEDODPLVAMGTMRLGAYPOLLKPGSKAHKAYNESS HGLR I VGTCPPOGLCE I I EVODHPMM I GVQFHPEFVSKI ANSLEMDPNTPHPIVY RHRYEVNPDYIOSLED PLFIAFIEAALVYSKDA

273741 274214

yggF Pamily SILRMOAMSKPSSCKAYLGIDYGKKRIGLAYAAEPLLLTLPIGNIEAGKNLKLSAEALHK LEGISTER STEELENGE GOODS, ECONOMIC SERVER STEELEN LEGISTERS STEELE STEEL

CPn_0238 274210 275838

zwf-Glucose-6-P Dehyrogenase
PCNIQKLRDFNRFILLFVIFASAGTKKEIKMTNVVQETIGGLNSPRTCPPCILVIFGAT
GDLTARKLLPALYHLTKEGRLSOP/CVGFARREKSNELFRQEMKQAVIQFSPSELDIKV
WEDFQQRLFYHRSEFDNNMGYTSLKDSLEDLDKTYGTRONRLFYLSTPPQYFSRIIBNLN
KHKLFYKNQDQKPWSRVIIEKPFGRDLDSAKQLQQCINENLNENSVYHIDHYLGKETVQ
NILTTRFANTIFESCWNSQYIDHVQISLSETIGIGSRGNFFEKSGMLRDMYQNNMMQLLC
LLTHEPPTTFDADEIRKEKIKLIQRISPFSEGSSIVRGQYGFGTVQGVSVLGYBEEENVD
KDSRVETYVALKTVINNPRWLGUPFYLTARGKELAKKSTDISIIFKKSPYNLFAAEECSRC
PIENDLLIIRIQPDEGVALKFNCKVPGTNNIVRPVKMDFRYDSYFGTTTPEAYERLLCDC
IIGDRTLFTCGDEVMASWKLFTPVLEEWDQDSSPSFPNYPAGSSGPKEADALIERDGRSW
RPL

CPn_0239 275863 276672
devB-Glucose-6-P Dehyrogenase (DevB family)
KSISMTNIGIETMATLINFNDTNKLLLTKOPSLFIDLASKDWIASANQAIKQRGAFYVAL
SGCKTPLEIYKDIVINKDKLIDPSKIFLFWGDERLAPITSSESNYGQAMSILRDLNIPDE
OIFFMETENPOGAKKYQELIENKIPDASFDMIMLGJEDGHTLSLFSNTSALEEENDLVV
FNSVPHLETERMTLTFPCVHKGKHVVVYVQGENKKPILKSVFFSEGREEKLYPIERVGRD
RSPLFWIISPESYDIADFDNISSIYKMDIL

277861 276698 No robust homolog present in Genebank/EMBL as of 11/7/98
LVYFMVFSPSSESVVKANSVVRSNFCYFLENKFVSPSESTEVMFSEIMKGRVPDIESLFD
RPTDMMTGFKAAQNLGNLFNSFGILIMCFSQCKSCQTPEKETSAIVLGATLLFFVVALI RPTIMMMIG-KAAQNIGNENSFGILIME-SQCKSCYPERETSATVIGATLLFFVVALI
LGPTIGALVYCAYKVYTIGKMIYSLIMKAKAKVLRHPAQNYFHRAACVATIRSSEEAVKAC
KLYKSAMIGSLVVSLIASLALIALTAGIVLVLFFVAPGAAPVITAAMMGCCAAGGGALLI
SLIGLMIAIVRKAKHQEACVGHLTNVVLHTAVSEALLHDPSHFQTNALARDLFLTDCLSH
YGHLFSNEEVAQLVQGGAPGGGSRPSQHYGGSSDYQNRRGGNGNFGGSHFGGGGFAGSH
FGÄGYPTAPTMPSAPPPFPPPAYDTIYG

279372 278203 CPHED241 279372 278203

NO PODUST homolog present in Genebank/EMBL as of 11/7/98

IFEMREMSAMISLSSSHEASIASHTOVROVLVSLAMDEFVENNTEILPIKVFLARGTLSS

TAIJDLKDVVETEGEHHFQVYSNISLKMIYQRFFEKIFGIGCCPLLLVTDSHHTDPCGA

LITGIFAAVLFTVLAIVFOFTLGILCYSAYKIYQLTKKISSLSRTHTEVINSVQKSDFFI
HREĞAVAAAASQSTIKACKVFRQSTLIFFVLGLIITISLAALIVGLVFALFFLDPGAPA

WHTAAMIGCCAAGGTGILLSVIGFLLASVYSVQKSQEGVHHMHTALLRCIVSNTIIOMPY
LPJEFGTKKVLTOSIRRYQOFFSDDEYRDIESEVPLNRQTTPPPSYETLFHEEGSDGSSN VIPRESPPAYSTIDSSNSPFPSSSPPPYYR

CPD_B0242 279975 279487
NO FIGURE 10 CONTROL OF THE CONTROL OF THE

CPn_0243 280609 280133
No robust homolog present in Genebank/EMBL as of 11/7/98
INNEMYLYFLLKFVKGRIIMACSIGYHLCNANEPDRFVASKVALVADILLYPFMAVICAVV
FAYLMVKLLFLAIKFLVNTCIAACKSRPLPSCKENFOCLFGPKDKFOPSDMLGCLVLIF IIGTLIYSTIITVQSDTNRLRYFIISPAYQVGSTAIINW

280906 281556 CPHEU244 280906 281556
adk_adenylate Kinase
GAPSYTKGSVFIIMGPPGSGKGTQSQYLANRIGLPHISTGDLLRAIIREGTPNGLKAKAY
LDKGAFVPSDFVMEILKEKLQSQACSKGCIIDGFPRTLDQAHLLDSFLMDVHSNYTVIFL
EISEDEILKRVCSRFLCPSCSRIYNTSQCHTECPDCHVPLIRRSDDTPEIIKERUTKYQE
RTAPVIAYYDSLGKLCRVSSENKEDLVFEDILKCIYK

281627 282499 CPH_U245
ydhO-Polysaccharide Hydrolase-Invasin Repeat Family
TCOKEIMKHYLSFSPSADFFSKQGAIETQVLFGERVLVKGSTCYAYSQLFHWELLWKPYP
GHSFRSTLVPCTEFFHIHPNVSVVSVDAFLDFWGIPLFPGTLLHVNSQNTYIFPKDILNH
WNTIWGSGTPQCDPRHLRRLNYNFFAELLIKDADLLLNFPYWGGRSVH9ELEKFGVDCS
GFINILVQAGCYMYPRNAADQYADCHWISSFENLPSGGLIFLYPKEEKRYISHVMLKQDSS TLIHASGGGKKVEYFILEQDGKFLDSTYLFFRNNQRGRAFFGIPRKRK

282955 282551 rsy-sy Ribosomal Protein VVAKSTIQESVATGRRKQAVSSVRLRPGSGKIDVNGKSFEDYFPLFIQRTTILSPLKKIT EDOSQYDLIIRVSGGIQGOVIATRLGLARALLKENEENRQDLK9CGFLTRDPRKKERKK YGHKKARKSFQFSKR

CPn_0247 283430 282969
r113-L13 Ribosomal Protein
DDYTIMEKRKDTKTTIVKSSETTKSMYVVDAAGKTLGPLSSEVAKILRGKHKVTYTPHVA
MUDIVIVINAEKVRLTGAKKOGKIYRYYTGYISGMEEIPFINMMARKPNYIIEHAIKGMM
PROTELGKKQLKULRIVKGDSYETFESOKFILLDI

283650 CPI_0248 284453 283650
yetVybba ABI Trungdotter ATTUSE
ROPEOYLERWRYATEURUSEROPACKESRRIAGERISFORLAMGELIEAKNESKTIQQON
ON LID LIDVISCULIA GET IS TRUSCONEKTELLHELD TEDVPEGGGLAFFENDEKNOBLA
NERBALI GEVFORMELLEDDTVEKINIMEL LARKI ISKOSPYYTRALELEDLIVLEDKY
REBCHESKESKESKORVATABALINEFATELADEPSKILDEFTSEDTHNELLEQASALCGTE
LYTINKHLADFSCHASYJENIKLEFFINE

285850 CP15t hypotherical profein LEVMKFEFSVALKYLIPGROUDS IVSLF7/GIISLVVWLSIVFISVIHGLEQWIEDE SQLHSPITILPSDTYYSSYY HSSLS/FITTKTLOEKIASPQVDPYDPESDYLLPET FPLKDCOLGCQOKDPVMTLESSGPYLOSØHGKVIEFEGGVDYLDIKTSLKLÖKRQPRNL THFLTYPSKLSYEDKVLPYDETDYTSAELNPFNRSPSGWQQDFHLLEELYRGASIILPST YKDSGYKVGDTGVFSTYSIENEKETQYTYHVIGFYNPGLSPLGGRTVFIDPDLARSIRSQ SEGLGMSNGFHLFFPNTKRIVFVKKQIENILTSLGVDDYWEISSLHODYPOPILDQLQS DQVLFLFVCILILIVACSNIVTMSHLVMNKKKEIGILKAMTSSRSLKIIFACCGAFSG ACGVVIGTIFAIITLKNLOFIVKALNYLQGRETFNTAFFGQNLPNSVHPQAIYFLGLGTL

CPn_0250 25500; LS5502 rl33-L33 Ribosomal Protein KDSSMASKNREIIKLKSSESSDMYWTVKNKRKTTGRLELKKYDRKLRRHVIFKEAR

286036 287559 CMT_US1 289038 287339

**Conserved hypothetical protein
SPDSCLPWMSPFKKIVNRLICYISFOKESRTLPIIIREPRMTTKSLGSFNSVISKNKIHF
ISLGCSRNLVDSEVMLGIILKAGYESTNEIEDADYLILMTCAFLKSARDEAKDYLDHLID
VKKENAKIUTGCMTSNIKBELKPMSHIHTLLGSGDVENILSALESRESGEKISAKSYI
EMGEVPRQLSTPKHYAYLKVAEGCRKRCAFCIIPSIKGKLRSKPLDQILKEFRILVNKSV EMGEVPROLITERATATEARAGE.KRIKATETTETERKORLKSKELDILLEE KILVINSV KEILLIADOLEDVIKOVISTORSSOLESLLHELLKEERGOTWELRHLYLYPDEVSDGIIDLMO SNPKLLPYVDIPLOHIADRILKOMRRTTSREQILGFLEKLRAKVPOVYIRSSVIVGFFOE TQEEFGELADFIGGATIONLGIFLKSQEANTPAAELPDQIPEKVKESKLKILSQIQKRNV DKHNOKLIGEKIEAYIDNYHPETNILLLTARFYGQAPEVDPCIIVNEAKLVSHFGERCFIE ITGTAGYDLVGRVYKKSQNQALLKTSKA

CPn_0252 288112 287576
CT144 hypothetical protein (frame-shift with 0253?)
ATSTVCALWILDTYOSHDDAASCSFRACRFORYWLGGVNVPWNRFNOTSTOSTVINSAI
YIDSSOTWHNFROASAS IPRLFRISIFHTKHGDW IDNOTTGELLLVAYEANONPLFPDIR
IELAMSTCSGTSYYRARPMOWLCSTYYAVRPGYFVLENRSYSFRVOSFSWNIATLPFVN

CPn_0253/ 288474 287950
CT144 hypothetical protein (frame-shift with 0253?)
FCGGRUMSSSIPTTOKITISIPTFYRFILESINLTDEOKKTALTIGORIATENTOVLGNF
VDADGGLICONDLSVGGNINITPOTFNTMVFSGRVNLSNSPFSYQDSLGNKRQDYANINT EQPQQVVPYGYYKLTRVMMMQRAALSGGHVGSGDIGWGESMYLGISSIKRQHKVQ

CPn_0254 289268 288459.
CT1,43 hypothetical protein
IPMKTLGVKDONLFIDOATLSVERNVRIENNLETRDLKVLDTTTSPCEFIVKGNVSAEGS
QLANATTLSDEFITYSKTDVSGOFPVCNNISDPOSARDALTFSYYRKTCQAANLYTYYPGN
GYYVAPMITETHVAALITSKSVSRNATPDFSRYADIEPVVKLKQVGIYQVTMQLTRNSGO
VDGDNSATLILNFVSGNAKTLLCTSDTRGGYSSDRTSVAVTAIFSVTELVSSPPYDYPWI
NLESTIWMNLMSLSTCVIWFPFPSNFVEVD

290183 289329 29329
CT14) hypothetical protein
TLLKVIMKNNINNNECYFKLDSTVDGDLLAANLKTFDTQAQGISSTETFSVQGNATFKDQ
VSANGLISGTTYNLNAQNTISSQISIDFKNNRLSNCALPKEDCDEVPANYVRSPEYFFCS
KELIGDFDFNSGESYLPLTGSEFYLYQSRNVNSIFFIIGWKGSTRELTVGGNTAIDFLAA
CTYYVSFTVGKRWGWNNGWGGAIYINNGLGQVQCESTIYSGGGYATIGTLGTSIYRASVD VAPNPNDPNASDRYRAGIFYLSNGGSSAGIGNYSFSLLYYPDDRG

CPL_0256 291282 290398
CT144 hypothetical protein
FCGGRLMSNPFPKTKISIPTFVRFNIQSINLTEDQKKTTFTVGGKVTTENTVVRGDLTCT
DGGLTCQSDLT1QKNINIRPTSTNSMVTPGRLNLSNSPLSYKNSQGQDITDYEKMSSGKP
QEYVPFGYYKRTOIMMAQRAAHSSGYVGGGSVPSGYVPMNKFDQTSTQKTSGTEIYIDP
NDSTKLVFEVNNKVPKLFRISVIMAKHGSWLDNGTGADILLAANFYEGGGRINVTDLAM
TTSRGSSYYETRPLQVVCVTYYAQNNGYFTFQNRAGGGLRVSFFSWNIVALPYVE

CPn_0257 292136 291267
CT143 hypothetical protein
GVVMKRRNLQKILPNASTPSTRVAENTGIKDQNLFLDQATLNVDGNVDIENFLETRDLKV
ADTITSPCEFTVGGGLSAESSQFKATTLSKGLEITSEDQDGRVPKFTNVSDPQSPRDALT
YNYYRNTGCQALMLYTYYSSSOPTTVGKPIETVCQNPNPETYRISASAKIYDAVTRFPYI
OFKAPGIYQVTIQIRRESGGHSGLDNPNLYLNLMIGNNKTLLCASDTRGYSGGHRTSIAV TGTFTLTEIVATPPHDYPWLFLETTIGLDIKSMSTCVIWFPFQANFAEVD

CPn_0258 292534 292133
CT142 hypothetical protein (frame-shift with 0259?)
CFSFCRLGSKFEKITLGGNTAIQLLAAGTYILTFTIGKRWGWNNGWGGSIRLFEGKYTGD GTMLCGSTVYSGGGYSTIGYLSTAVYRDHSDIDPDPNNPSDKYMNNFLFVRNGDHSAVIG

CPn_0259 293031 292441 CT142 hypothetical protein (frame-shift with 0259?) LYPVFRKTYNYFIEMTTTNNODNNECYFKLDSTVDGDLLASNIOTFDKQAKGISSTETF SVOGNATFKEKVSATGLTSASTYKLNATGPAPSSITIOMKNNRLSNPALPRNPCDPVPAN YVRSPQYFFCAKPIEGTFMFDGSSRYLPITGDGSNYTLYQSSKAGDVFRFVDWDQNSKKL HLCGTOPYNFLLQEPIS

CPn_0260 294090 293548
seca-Protein Translocase Subunit
a/LDFSKRSCVEEDHVSKKINRIDLC-PCGSNKKYKQCCLKKEEQTARYTTEGKFKFSAEV
LSASEQGEAGONCTKLFORLG9SLTSEGKAAVGKFHQITKNIKEVMSKKALKKAQAKEEKL
VTEKLQOHNFEILNTGENLAPPMESTATLNQDTNFVCEDFIPTOEDFRISENSOKPPVEE

CPn. 0261 294272 CPILO261 294272 29513
yddO-PP-Loop Supertamily Affase
YSCHHER IVRMCTLLLINPFWMKACKRIESELVKALYTHTMLANHRKIVVALSCOKDSLTL
LLMLKALSCREPDLDLHAVNICSKYSCSAEVNKPYLTRICDQLCIPRRIPSPYAPETP
ECYTCOGARRELLFDANKEIGASATAFSHHRDDLVOTALLHLLHKAFFASHLPVLDMVHF
GYTILREL IFTPEFWIRKFAKEINFARVTCRCPVVSLRSVAEQSLKLLEEVFPLARHNIA 295013

CPn_0262 295055, 295033 DBLE:GBLE:Tike Acid Thoophatasa LIPNINKEVKYYEVLMNKREKTIETNDESTTAKOMBOLVSALLEANIGDTYTAAPQAEQG

HEIMALULNOVVCASPYAYPOPVKEAWAVGGSPTDCVT LFESVSPDLV NN ICKNAWYGIT IGAAKQALVDG IPGMALSQONNISF PELIKALVITY CLITTILN IN PTSPGGSSWECHRLVPPTDEFFYEEPOYL NOVYVGKISG GEELACMLENHIGVSPIFSONSPIGLMTLEEFOKTQENFNASLLSSELTTKIF LFESVSPDLVISGINCG PEILKALVIYLLSOPFP KNOYYVGKISGVRIGEHP

296174 297136

CYP_USB 2 2007 2 MODATE BELLE COLON DESCRIBERABIER DAVIENDE PVELARES FORME<mark>T FOMETA</mark>N ET COLONIA DAVIER WEIMORE VERSCHMET VERFALL GVOCCUL FRICKET DOTE LLO IIINKKKGYTVOOIILFVNFFIFALSGIVYKNWHTAFVSFLTYGIATKVMDMVILGLEDT KSVTIITSSPRKLGHILMETLGIGLTYIHAEGGYSGEPRNLLYVVVERLQLSQLKEIVHR

EDPSAFIAIENLHEVINGRRT

297730 297155 wbiD-Phenylacrylate Decarboxylase
MKRYVVGISGASGVILAVKLIKELVNAKHQVEVIISPSGRKTLYYELGCQSFDALFSEEN LEYIHTHSIQAIESSLASGSCPVEATIIIPCSMTTVAAISIGLADNLLRRVADVALKERR PLILVPRETPLHTIHLENLLKLSKSGATIFPPMPMWYFKPQSVEDLENALVGKILAYLNI PSDLTKQWSNPE

CPn_0265 298632 . 297730
ubia-Benzoate Octaphenyltransferase
KIIIVRLNYFLNLUNFKYSIFSILFLSASTVFALSINEISONLSFKEGFKISVFGAIAFV
FARTTGIVVNOCIDRFIDKKNTRTSKRVLPANLVSLNFAWVLSLFCSFLFLFLCKILRIF
SLGIASLTLMIVYPYMKRVTFFCHWGLGLVYTVAILMMFCAFAESGLSMRLCFLALLWGG
SVGMVIAANDIIYAIEDTEFDREEGLRSVPAHYGEKKAVEIAKVNLWVSYLAYIFSGFVG SLDKEFYFTAIIPLVVILKVVRMYSNYSKKDQEGESKFFLANIAIALSFLVSMTLFWSLS

CPn_0266 299181 299876
No robust homolog present in Genebank/EMBL as of 11/7/98
IMALDEINNONNESQOIASSTSQTSKINQDRRTFACTVYLLVVARLMILSGIVLLFTIGS
LGLSVPLSGILGTFAVTVGAVLFITGLITILVRKSLGIEQKNEDLMFLKIKTPTPPARPLM
SKFSVTCSTTSIVLGMALLIGAVVSVFFLTGYLQLGLCAGLVGLGTALFVAGLARMSPRS
LADQEGSGSADSQSNIVGIGEPKAAQEQKWYKMAVVRGEDGIPTAIRLTPEK

CPn_0267 300122 300910 No robust homolog present in Genebank/EMBL as of 11/7/98

VSIMSLNKTNALLNOPEPAVCLNAWDPKYINODRKTFACTVTLLVIATLMILITGVIVLL AMGSPGLSVLVSTIIGTSVTTLGTALFIIGLVKLIKKSLAWIQYQKYFQEVVKQKYEPFS IPRNDNYHKLTSCLPSPLDIESPSPEASTPVSKLRIACSGVAIVLGVTLLIGAVVSVFFC
TCHEQLALCVGFACLGTALFVGGLAGLRTHSLIAQGIMYLYLTYYLSSALEERNETVKDQ RNEINTYLTEECROOKREKALLE

CRD_50268 300914 301318
NO.TODUST homolog present in Genebank/EMBL as of 11/7/98
KOMALSIMSCCOSSSTSTWEWMKSFVPNWKNFTPPLSPIPSEDEFILAYEPFVLPKTDPE
NAQANPPGTSTPNVENGIDDLNPLLGQPNEQNNANNPGTSGSNPTSLPAPERLPETEENS QEEEQGSQNNEDLIG

CP1-0269 302468 301476

Dipeptidase

VARROVATIOMHCDLLSHPHFCRKDPAVRCSPEQLLSGGVRQQVCAIFVPHSRGEPNCDK
QNSLFFSLPNQYPDIGLLSYEEEENGSSSQKKSLSLIRSIENASALGDDTAPLGTLLAKL
HLTKQGPLAYLGIVMKGDNRFGGGTEAPKRLSNDGKVLLDIMYELGVPIDLSHCSKKLA
EDILDYTADKLPNLAVIASHSNFRSVLDHRRNLVDAHAKEIVRRKGVIGLNLVRSYVGDS
LGDLEKHVLHAENLGILSSIVLGSDFFYANDEDENFFFNECSSAEAHPVLNQLIHRIFSKG KAESILSSRAEKFLKQVIVEQVNPKITDVKL

CPE-0270 303343 302468
ywlC-SuA5 Superfamily-related Protein
SIFGVIVPDKKAGITFSLPEVMSAIHQGKIVALPTDTVYGFVLSLYASEAEERLYALKDR
EPSKAFALYVNSIEDIENISGYPLSPTAKKLAQLFFGAITLVVKHRNPRFPKETLAFRIV
DHSWREIVDHCOTLIGTSANLSEFPSALTAQEIFADFADHDLIFDGPSKGLESTVVA
SDPLYIYREGLISRSVIENIAGTEAKIFHRTSHAFSKHIKIYTVKNQEQLVSFLSGSLDF
KGVWCEHPKPKNFYTRLREALKKKTPSIVFIYDINTSDYPELFPFLSPYYIE

CPn_0271 303628 304362
Lysophospholipase esterase
KLMTDYSFFRRKIGNIEAIECFONPODPIIILCHGYGGLADNLTFFPSICSFSKLRPTWI
FPNGILPLENDFRGSRACFPLNVLLIQELSRLYANGVGNLQEKYDELFDVDLETPKEALE
ELILNLNRPYNEIIIGGFSQGAILATHLVLTSGNPYAGALIFAGARLFNQCMEEGLKQCA
QVPFLQSHGYEDEILPYHLGAHLNDLLTKLNGQFVSFHGGHEIPSVVFQKMQVTVPNWI

CPn_0272 305272 304340
dnaX-DNA POl III Gamma and Tau
FNRQSDATYATWYMHLEEENQGWEALLRKYYHQEVPPAILLHGFTLPVLQDKAEQLASEI
LLSSSPGSEHKYSQKIHPDIYQFFPEGKGRLHSIDLPRGIKKQIYISPFEANYKIYIIHP
ADRMTLAAISAFLKVFEEPPKHAVIILTTAKVQRLPKTIISRSLSIFIERGEKILGSKET
FSYLFRYAQCEIPVTEVSQIIKESSETDKQVLRDKVQRFMEVLLELYRDRYTLNIGLKJKS ALNYPEHVKEILOLPLLPLDKVLLIVESACRSLNNSSSAASVLEWVAIQLVSLQYKEKEL VSVSPGODLSN

CPn_0273 305853 305227

CPG_0273 30593 305227

tdk-Thymidylate Kinase
GGIVFTVIEGGEGGGKSSLAKALGDQLVAQDRKVLLTREPGGCLIGERLRDLILEPPHLE
LSPCCELFLFLGGRAQHIQEVIIPALRDGYIVICERFHDSTIVYQGIAEGLGADFVADLC
SKVVQFTPFLPNFVLLLDIPADIGLQRKHRQKVFDKFEKKPLSYHNRIREGFLSLASADP BRYLVLDARESLASLIDKVMLHTOLGLCT

308368 305852

TYTA DNA GYTTGE SUBURIT A ESTTIEMFUKDETIVEKNUEEEMKESYLRYSMSVIISPALPDIRDJLKPYORRVLYAMKOL CLOSCARHERCAR FCOTECOPHICIES TYPPLE PROMOTES AND TERRAPOTER CHESTORY VANAROL

FLAME TEARLY III AMYLMEDLORD TV POLY PROMOTES TO A

FLAME TEARLY III AMYLMEDLORD TV POLY POLY PROMOTES TO A

FRAME TEARLY III AMYLMEDLORD TV POLY POLY POLY POLY PROMOTES TO COCCOGINS AVT

GRICK I REPART LIVERING BORDER FOR TITEMPYNNIK OPLI EQIANLY NEKTLAGIS DEVIDE

CLARLY PARLIMER FOR THE MY REPOLY VERTAM LADRIL PREMIS HARISANI

RHREET I PRITTY ELAKAETRAH VLEGYLKALSCLOALVKTIRE GONKEHAKERI I ESFO TTEPOALAILELRUYOLTO CAEYEELLNKIAYYKOVLSDEGUYKDI IRNEODL LKHHKVARRTTIEFDADDIN LITNEGVI ITISGDDYVKRMPVKYFKEORROGHGVT GFDMKKGAGFLKAVYGAFTKDI LLITTNEGOCYMLKVMOLPEGERBAKGKPI INFLEGIR GEELAAILMI KNFONAGFLELATKGVVKKVSLDAFSNRKKGI RALE IDEGDEL TAAC HIVSDEEKVALFTHLGMAVFFHEKVRFMGKTARGVRGVSLKNEEDDKVVSCOJ UTENGSV LIVCDQGFCKRSLVEDFRETNRGGVGVRSI LINERNGNVLGA I PVTDHDSILLMSSQQA IRINMODVRYMGRSTYGYRLYALKEGDALVSMEKLSSNENDDBYLSGSEEECSDTVSLR

310 744

PYTE TNA CYCLE COMERCE BY THE TRANSPORT OF THE TRANSPORT MGEEVPPRREFIESHALSIRINNLDI

3/0793

CTI91 hypothetical protein /
DMFLKRKKRGGSQVQNKRTASPIKHAKYYLHNYLQELQKIMAARPHDAIDAWNQVFRDKY
KGMSQAIGFRDHILLVKVYNSSLYALKKQTPQNDLIMSLYQVASHVQIREIQFLLG

CPn_0277 312003 311404
No robust homolog present in Genebank/EMBL as of 11/7/98
NISIFYPKYFIEGKEVLIKNLPPLFYGVILMIINVRAPAFGITSVQOFSTNFQAAIPIL
NIVIGCSRISSTYAEDIEEVAQBKLEKSTHSKSSTSVNLWAHRVRGVVEILGGGIVILAL
EITALVLQVIIKLIKCLIDVLC/CLFGLGVCVVAIIGAIAFCVVVVVKYLGFCSQGEELE
PIEVKTLISPDKPYPTVVVV

31/2884 312060 CPT_U278

37884

31200

**conserved outer membrane lipoprotein

RDSMKKKLSLLVGLIFVLSSCHKEDAQNKIRIVASPTPHAELLESLQEEAKDLGIKLKIL

FVDDYRIPNRLLLDKQVDANYFQHQAFLDDECERYDCKGELVVIAKVALEPQAIYSKKHS

SLERLKSQKKLTIAIPVDRTNAQRALHLLEECGLIVCKGPANLMTAKDVGKENRSINI

LEVSAPLLVGSLPDVDAAVIPGNFAIAANLSPKKDSLCLEDLSVSKYTNLVVIRSEDVGS

PKMIKLQKLFQSPSVHFFDTKYHGNILTMTQDNG

CPn_0279 313546 312875

* Possible ABC Transporter Permease Protein
KKIMGSPLIGILIKETVATLYMVSTAFFFSCAIGGMLGIGLECTSPKSLNPKKSLYATIS
MILSFLAIPPALLIVLIFPITRMIVCTSLOFTASIVPLIGAIPPVVTIVVDAFRNSAL
NYLESAVALGIPKNNILFGILLPESYPOLIFSLKSLVVHLISCSTLAGFVGGGGLGOLLL
QYGYYRROSTTSVIDVITLVLIESVRILGDFWGRRVLKYRGIL

CPn_028V 314593 313550
dppF-Dirdeptide/Transporter ATPase
IKGEAMU STOHSPIISVODVSKKLGDHILLSKVSFSVYPGEVFGIVGHSGSGKTTLLRC
LDFLDMPTSGSIGVAGFDNSLPTOKFSRRNSKKVAYISONYGLFSSKTVFENIAYPLRI
HHSEMSISSEVJEOVYDTINFILLYHRIDAYPGNISGGOKGVAVIARAIVCOPEVVLCDEI
TSALIPKSTENIIERLIOLNOERGITLVLVSHEIDVVKKICSHVLVMHCGAVEELGTTEE
LFFLNSENSITNELFHEDINIALSSCYFAEDREEVLRINFSKELAIGGIISKVIQTIGLVS
INLSGRINDFRKSPMGFLIIVLEGEVPGRKKAKELLIELGVIKEFY INILSGNINLFRKSPMGFLIIVLEGEVEQRKKAKELLIELGVVIKEFY

CPn_0281 ∮n_0281 315033 316103 dhnA-Predicted 1.6-Fructose Biphosphate Aldolase (dehydrin

CAMILY) ISLRRHTIMLNIHDILGNDDENLLSYQCKHITKDKLTLPSHDFVDKVFGLSDRNNRVLRS
LQTMFSHGRLANSGYLSILPVDQGIEHSAGASFAINPIYFDPENIVKLAIESGCSAVAST
YGTLSLLSRKYAHKIFFMLKLHNELLSYPTKYHQIFFTQVEAAYSMGAVAVGATVYFGS
ETSNEEIVAVSNAFAKARSLGLATVLWCYLRNPAFVANGVDYHTAADLTGQADHLGATLG
ADIVKQKLPTCQGGFKAINFGKTDERVYSELSSHNPIDLCRYQVLNSYCGKVGLINSGGP
SGKNDFTEAARTAVINKRAGGMGLILGRKAFQRPLSEGIQLLNLVQDIYLDPNITIA

CPn_0282 316084 317529

xasA/gadC-Amino Acid Transporter
ILILOSINFSKKVFHHSHSKPTKPLGTFTVGMLSLAVVISLRNLPLTAKHGLSTLFFYGL
AVICFMIPYALISAELASFKPQGIYIWARDALGKWMGFFAIWMOWFHNMTWYPAVLAFIA
STIVYKINPELAHNKYIATVILAGFWILTFFNFLGITSSALFSSICVIIGTLIFGVILV
SLALFWIFSGNPIAISLSWGNLLPNFSNVSSLVLLAGMLLALCGLEANANLASDMVNPRK
NYPKAVFIGAIATLTILVLGSLSIAIVIPKEEISLVSGL/KTFTLFFDKYNLSWMTGIVV
VMTIAGSLGELNAWMFACTKGLFISTQNDCLPRLFKKVISKNVPTNLMLFGGIVVTIFTL
LFILCDSADLVYWILTALSVGMYLAMYICLFLAGPILRIKEPRAQRLYSVPGKFLGICTM
SILGILSCAFALWVSFLPPRELAQISEGSKIGYTTFLLLAFSLNCLIPFGIYFTHKRLSK
KS

CPn_0283 318581 317532
No robust homolog present in Genebank/EMBL as of 11/7/98
GRRLGYFQDLIKNAVAKIISFRKSPPNEVKLLIKFAKKDLKNSSIAFLYEVLLEILEAPG
EFILEVLFSLDPMULKSMLDPKKHSTLGIEISSFRAETIESSSLGLISINLLISGLCLRS
SHORGOAVKIIQOFCPGFSSEEVONFVEQRNILTPFLHHLFEGDEVALLNQLSURLDLIV
PNALYZEPPDSCWGINSERCAKDAEDQGEDFNKTKEACKEGLKKLVLPALGITSIFQLL
RARRFKQGAEILMAITRKKMKONFFIFLEALLESEEFGIGVKYLKLIMFIHLWDKLLHA
LYLGYFOYGLIGOGELETFGRRANINPEAFOAAIOQOGLLGFLFFKMLLD

119054 318551 CPT_0284 19054 31851
No cobect homolest present in Genebank/EMBL as of 1177708
FLIMHTPAPQVPVTDEPVNNTDSSSGLERSDERFETZLILATIA FATLMSVLYPGGTIG
VSTPVISMLTFLSVGSVLVAVLFYQQGGTEKTKVFGTTGPGVFFGDEPLNLLLGREEDG
VSATDELLKNFPADPERREMMLYSSHFLDEQGRPNEGFEEDGHTGKTL

CPH_9285 COUNT 319051
No robust homotom present in Genebank/EMPL as of 1177796
KELFPILFFTANKETASHELIYEKHOGESLSP/TIL/LLATSVELLES/VVFALWGCHVL

AAPLGLLVWGCAASVCSMMAIVSLMCLYKGGKPLIEPS
PVEQGSLPKERKTVSFKAKIPSIVEDDFKPYVIGSTF
TLIVDFPRALEESSKSSGSLLRGVISEIKNLFLPRFLGS
SLTACLPRLGSIVEEYA
SSDLIILLIKREPELMNYTQOLIAHLINSLKTEKRKLTPHMOKUVLSINFWFYGWISEEKC
IEKIVAYDPNLLTDELKAHLEAGNIVOFLLSFOSSEMGREFRALFPSDAQELPSAKDGSN
YVPAINSSEYMFTSSVASHYKDFSLLFTFF
SNQSVILGNPFLLIELLHENPKCOTFLKGLLEKAMPMSNWAALFRPMLMGMLCSGIARK
KELKIIAEHLTVPFKEITQAIASGKILDLLLQHLFDF

mgtE-Mg++ Transporter (CBS Domain) 320650 mgle-Mg++ Transporter (CBS Domain)
SCRESKGKIMVGEONRHEEKLDTAFSSGNIMDSRTSHLDDELSFKLEKAFTCLSTDIHSH
DLSKIVIETNPIDLAYAVSCLPSESRAILYKNLSCITAKVAFIINTDSASRWAIFRRLSD
SEVCALIEQMPPDEAVWVLDDIPDRRYRRILELIDSKKALKIRDLQKHGRNTAGRLHTNE
FFAFLMETTVKDVSACIRSNPGIDLTRLVFVLDFKGELQGVVTDRSLIINPPEMSLKQIM
NQIEHKVLPDATREEVVDLVERYKIAALPVVDEENFLIGAITYEDVVEAIEDIADETIAR
MAGTTEDVGYQTCHVVQRFLLRAPWLLVTLFAGLISASVMAYFQKISPALLALIIFFIPL
INGMSGNVGVQCSTILVRSMATGTLSFGRRRETIFKEMSIGLITGVVLGILCGLVVYLMG
FLGLNIFSGGGIGLGVTVATGVVGASLTATTLGVLSPFFFAKLGVDPALASGPIVTALND
IMSMIJFFLIAGGINFLIFFN

IMSMI IFFLIAGGINFLFFN

CPn_0287
324230
322089
No robust homolog present in Genebank/EMBL as of 11/7/98
RRCMIRSPLPFISSKRALNMLGLQDEFSCPEDVVDFLFSEIELLAQQDEPSEGYLALSRS
LLMMTHNHPKVVKRVIFYGVSYGLKHKSMSIFIDVLTYIDFLFEKLGISASDRLSLCSAR
TCINFELYSQYGEMKELSEVVDNFRLIEGLLKHHPDLKNRLGMEHFRIGAKQEEVSLVAS
ASVYQAVGRSFIELYHKHLELSDLACGMKCLALALDLSPNNAHIHADYAKGLVVLGTRQG
KSLLIERGMEHFSKAIFLSFSRDGDTLAYQNYRYSYALASVKLFDLTYKKEHFDQAMNIL
YQTVQAFPNLSGLMWWGELIRSGKMLNSNMMYIEVGLEFLASLQKKTNDPIALSGLAT
GIAILGLYLEEPNLFKDSRHRLISAMRTFPGNSALVHALGVVQLCSALYFNEDSHFASAI
SCFQSCLEWDLDATGMWQKLFDAYFSWGIKKKSARLLRKAVDVASRLCSLRPEAFLFWSD
RGLALKCLAEATIDEAYKEIFLSESLHYVQRAWDLSGRLEILELWGQSHYLLAELQGSLF
HYDEAYTLLTKVDLTLSSSRVKLILAAVLLGKGRLLQDTDPAEEAREILEPLVEVYLEDE
NFLLLLGKVYLFLFWKNKNVCLGKLARTYLEKATSLGCPEAYYTLGKFYAVIKDVMKAWG
MVIRSAQYGVRITEAKWLNDPYLANLREIHAFREVVENQKGRLWLGNKTEMKRN

325785 324571 CT288 hypothetical protein ISITIREFLFFGFECRAKFYNVIMSCFNLTSTNESLRPISPKASFPKQGWQSYFRSALRK HSDTLSVSVCKVNKYDANLFVRLTVIALAVVGVLILFSIMLASIQGTLVITSWPLVTRA
LĒLPTILLTOGMYILHRLGKKVDVISGVCIPPFSRRCWVPISSSHTLEKFDEKHVSACSY
LDISTLSADOSGIRAVYQCPPLLFRAFPCFGIPCAMPFVALLRMYNNLIRFLVVPFYIIF
RHĪVĒHFFCKHLPEDDRFTYKDVAREMGRSLAAFIKAPFYASACMIGAFYSLLDPLAGRV
LMGSVERDWNDNVILARSVSLANEAHSLFRFEGGGGRKGLGQHAFYLMLCCQPQSVFLFD
KGĒTVSGAHPSIQLPERRGLDTSGRYPHISVIPDSGNDSAKNFIV

325797 326996 CTIND hypothetical protein
NEWRIMKKORSHYKKNNLLLLLSILVGLGLGSVOSPWIVYSAECIANTFLKFLRLLSIPL NENGLARAGOSHI KRANDLLISIDUGISISYOSHU YYSACLANTILKYLKILISIPL VEČALGSTITSIONFNIMTILGKRILYYTLLITVIAASIGILIFFILRPOHITODALATT TKCAPPLGYLDVLSDTLPENIFKPFLOGNVISAACLAVLLGTASLFLQEKEKHFVNQFFNS FRŠIFINLARGGLKLLPIAMLGFSVILFKELKOSNILIFFAEYLLCVIGANLAGGFIVLP IĞEKINKVSPLKVAKAMSPALVTAFFSKSSAATLPLIMELAEDDLKINKNISRFSFPLCS VIRNACCAAFILITVLFVATSNGMIISPLMSLGWIFIATLAAIGNAGVPMGCYFLTLSLL TSMNVPLSILGLILPFYTVIDMIETSLNVWSDCCVVSLAN

327027 328523 Na-dependent Transporter
RSALTMNKKHASFSSRLGFIFSMIGIAVGAGNIWRFPRVAAQNGGGAFLILWLCFLFLWS IPLIIIELSIGKLTKKAPIGALIKTAGKKFAWAGGFITLVTTCILAYYSTIVGWGLSYF YAWSGKIHLGNDFAKLWTSHYQSSIPLWAHLTSLGLAYLVIRKGIVHGIEKCNKILIPA IMPOGRIFILD AALWISHIQSSIFLWARLITSLGLAYLVIRKGIVHGIEKCNKILIPAF FLCTIALLLRAVTLPGAVQSIKQLFSCDKSCFSNYKWIEALTONAWDTGAGMGILLLVA GFASKKTGVVSNGALTAICNNLVSLIMGIIIFSTCASLDILGTTQLQDGAGASSIGIFFI YLBELFTRLPGGIYLTTLFSSIFFLAFSMAALSSMISMIFLLSQTLAEFGIKPYISFILA TIŠAFVLGIPSALSLTFFSNQDTVAVALIVNGLIFIYAALVYGFPKLKKEVINAYFGGL RLNKAFDYIIKYLLPIEGILLLGWYFYEGLFPENGOWWNPISLYSLGSLVLOWSJCLIIL

CPn_0291 328658 329194
incB-Inclusion Membrane Protein B
EKHMSAPIPTPQELSDQITCLNVQYQOVSELARENKGDIEGLKTLTAALTADAGIQPSAD
EIYSLQTAAALILSASEKROSGPSGSTEGSVTVQSFCKFKKVLAVVLTIIJLIATAVLIA
CIIAACGGFPLLLSALNLYTIGACVSLPIIASTSVALICLCTFVANSLIYPVITVRTTR

329201 329836 CPD_0292

J29201

J29836

incC-Inclusion Membrane Protein C

VKNTKNSDFMTSPIPPQSSCDASFLAEQPQQLPSTSESQLVTQLLTMKHTQALSETVLQ

QQRRLPTASIILQVGGAPTQGAGAPFQPGPADDHHHPIPPPVVPAQIETEITTIRSELQ

LMRSTLQQSTKGARTGVLVVTAILMTISLAIIIIILAVLGFTGVJPQVALLMQGETNLI WAMVSGSIICFIALIGTLGLILTNKNTPLPAS

CPn_0293 329940 332723

CT234 hypothetical protein
WSMORVLRLENLHHGEEKRAFLFFLLGLVWGIGGYGTLSLAEGLFIEKLGSAELPKIY
LGSSLILGVLSSLILLVLNERKHHSATALFELIPVSLSILCNFYLILSSIFAIDPRSPLFF
YRIVIWSLTILSYTSFWGFVDQFFNLQDGKRHFCIFNAIJFLGDAIGSGIIASLVHTIGI
GGILLLFTAALLTFFIVYVSKSLKSLSDDHDLFIDTGHPPELSKALKLGFYDKYTFYL
LCFYFLWQLLAIATEFNYVLKIFEIOFASKEEFELVAHIJKKSLWISIGGNOFALFAYSYL
LCFYFLWQLLAIATEFNYVLKIFEIOFASKEEFELVAHIJKKSLWISIGGNOFALFAYSYL
VRRLGVNNIILFAPLAFLSLFLFWFKTTLSIAVLAWVREGVTYALDDNNLQLLIYGVP
NKIRWJRI IVVESFIEPIGMLVWSLICFLSSQOYVFGLIISLIATILVCLVRGYYAKAIL
KNLGAOALQLTRSMGDWIKOMTVKOKROVELFLLAHLKHPSERHOTFAPOHLLNLASRGV
LPSLLAHMNKLSLFNNLKTIEMVKSSLWAKDFLTLJLLKRWTSIFPHPAIAGAIHLYFAR
LDLLHITHITAEDLYDTVGDRLLAA LTVYRGGAVYDFNDLADKRLKELINSOPPEDIVAG
LTILKLEKNPOMFPILLDFLNTKNEDILITTGCAKTYDVRANHKPYCFELLKRLRCGSHN
DEAGOYLLKTIGIALDISPVKDLMTTEQLKNTTRYARAMIGELDKEVAPAFLOVLTDE
THENEGRICAAKALCKIONWLLKKHAYKIVKJSGASALFYSYGIGHYIOKKYPTYNLDLLA
NTIAENYYAAEVMFMIGILGILGSMEHSGVLIFALTSKNOKIKAQALESLEKIVJDGHLFGL
CLEFFYNORISMCYSERYYRKGOVIPTLEKELLMMENSPOSLNKUTAQQLKEELSYCDPDF
QSVNTIYNOGHEDFRTEESETLISFLSI

CPn_0294 333077 11/502

WKFNKQLYLRFSRYNHEIL

CAMP-Dependent Prot Communication Requilitory Subunit IRNFFMNLIDRAFLLKKTI MODILLYTIADKTETLIFKNOSNYFSIGOPGFSFYII VEGYITISKEKLESPLNLKPLD FGEESUFNNKPREYNASANTOVRMLVLGKGQILNIVE **ECPSVALSFLELYAKQIKFREP**

CPn_0295 333866 /333627 acpP-Acyl Carrier Protein AMSLEDDVIAIIVEQUADPREVNENSSFIEDLNADSLDLTELIMTLEEKFAFEISEEDA HELDTWINNETTERUPAED

334022 CPH_020b 33470£ 334022 CT296 hypothetical profein KIPIRGMICHDITLVGKKVIV7GSRGIGLGIVKLFLENGADVEIWGLNEERGQAVIESL TGLGGEVSFARVDVSHNGGVØCVQKFLDKHNKIDILVNNAGITRDNLLHRMSEDDVS ISTNLTSLYYTCSSVIRHMIKARSSIINVASIVAKIGSAGOTNYAAKAGIIAFTKSLA KEVAARNIRVNCLAPGFIEFDMTSVLNDNLKAEWLKSIPLGRAGTPEDVARVALFLASQL SSYMTAQTLVVDGGLTY

CPn_0297 335724 334774
fabD-Malonyl Acyy Carrier Transcyclase
SRSNKDDNFMKKRYAÐLFPGQGSQVVGNGQDLYMEYPEVRELFDFANERLGFSLTSIMFE
GPEDLLMETVHSQLAÝLYLHSMAVVKVLSQRSSIQPSLVSGLSIGETTALVASDRISVLDG
LELVRKRGQLMNEAKNQSPGAMAALLGLPSEVIEENTISLGGGIWIANYNAPKQLVVAGGI
AEKVDQAIELFRDJGCKKAVRLKVSGAFHTPLMQVAQDGLAPDIYALCMKDSSLPLVSHV
VGKSLVNTEENRÐLARQMTSPTLWYQSCYHIESEVDEFLELGPGKVLAGLNRSIGISKP
ITSLGTFAQIEKFLSEV

CPn_0298 336742 335717 CPn_0298 336742 335717
fabH-Oxoacy1 Carrier Protein Synthase III
YTSFFLYMWFSVNNNKKAAIMATGSYLPEKVLSNADLEKMVDTSDEWIVTRTGIKERRIA
GPQEYTSLMGAIAAEKAIANAGLSKDQIDCIIFSTAAPDYIFPSSGALAQAHLGIEDVPT
FDCQAACTGYLYGLSVAKAYVESGTYNNVLLIAADKLSSFVDYTDRNTCVLFGDGAACV
IGESRPGSLEINNLSLGADGKIGELLSLPAGGSRCPASKETLQSGKHFIAMEGKEVFKKV
VRRMETAKKHSIALAGIQEEDIDWFVPHQANERIIDALAKRFEIDESRVFKSVHKYGNTA
ASSVGTALDELVHTESIKLDDYLLLVAFGGGLSWGAVVLKQV

CPn_0299 336726 337415
reck-Recombination Protein
RKMLVYYSESLYSKINLGPRPECKNKIHITMTRYPDYLSKLIFFLRKLPGIGFKTAEKLA
FELISWDSEQLKILGAAFHNVASERSHCPLCFTLKESKEADCHFCREERDNOSLCIVASP
KDVFFLERSKVFKGRYHVLGSLLSPITCKHIENERLSILKSRIETLCPKEIILAIDATLE
GDATALFLKQELQHFSVNISRLALGLPIGLSFDYVDSGTLARAFSGRHSY

337768 340152 yaeT-Omp85 Analog GRLLGMLIMRNKVILQISILALIQTPLTLFSTEKVKEGHVVVDSITIITEGENASNKHPL GRLLCMLIMRNKVILQİSILALIQTPLTLFSTEKVKEGHVVVDSITIITEGENASNKHPL
PKLKTRSGALFSQLDFDEDLRILAKEYDSVEPKVEFSEKTNIALHLIAKPSIRNIHISG
RQVVPEHKILKTLQIYRNDLFEREKFLKGLDDLRTYYLKRGYFASSVDYSLENQEKGHI
DVLIKINGSPCGKIKQLTFSGISRSEKSDIQEFIQTKQHSTTTSWFTGAGLYHPDIVEQD
SLAITNYLHNGYADAIVNSHYDLDDKGNILLYMDIDRGSRYTLGHVHIQGFEVLPKRLI
EKQSQVGPNDLYCPBKINDGAHKIKQTTAKYGYINTNVDVLFIPHATRFIYDVTYEVSEG
SPYKVGLIKITGNTHTKSDVILHETSLFFGDTFNRLKLEDTEQRLRNTGYFQSVSVYTVR
SQLDPMGADQYRDIFVEVKETTGNILGIFLGFSLDNLFGG ELLSESNFDLFGARNIFS
KGFRCLRGGGEHLFLKANFGDKVTDYTLKWTKPHFLNTPWILGIELDKSINRALSKDYAV
QTYGGNYSTTYILWEHLKYGLFYRGSQTSLHEKRKFLLGPNIDSNKGFVSAAGVNLNYDS
VDSPRTPTTGIRGGYTFEVSGLGGTYHFTKLSLNSSIYRKLTRKGILKIK KGRADFIKPYS
NTTAEGVPVSERFFLGGETTVRGYKSFIIGPKYSATEPQGGLSSLLISEEFQYPLIRQPN
ISAFVFLDSGFVGLQEYXISLKDLRSSAGFGLRFDVMNNYPVMLGFGWPFRPTETLNGEK
IDDVSORFFFALGGMF IDVSORFFFALGGMF

CPn 0301 340163 340762 CFT_0301
(OMPH-Like Outer Membrane Protein)
IKOLSKEIFVVFRKGFWYPFSIFKLVQVIMKKLLFSTFLLVLGSTSAAHANLGYVNLKRC
LEESDLGKKETEELEAMKQOFVKNAEKIEEELTSIYNKLQDEDYMESLSDSASEELRKKF
EDLSGEYNAYOSOYYOSINOSNVKRIOKLIGEVKIAAESVRSKEKLEAILMEEAVLAIAP GTDKTTEI IAILNESFKKON

340766 341866 CPn_0302 340766 341866
lpxD-UDP Glucosamine N-Acyltransferase
skrkefsmseapvytlkQlaelLQVevQonietprisqvedisQaQphhiafldnekyssf
kkntkaGaiilsrsQamQhahlkknflitnespsltfqkcielfiepvtsgfpgihptav
ihptarieknvtiepyvvisqhahlgsdtyiGagsvigahsvuganclihpkvvirervl
mgnrvvvQpcavLoscofgyttnafedhkplkhldyvivQbdveiGahttjdkorfkriv
ihegtkidnqvQvahhveigkhsiivaQagiagstkigehviigggtgitghisiaDhvi
miaQtgytksitspgiyggaparpyQethrliakirnlpkteerlsklekQvpdlstpsl
afibsgi

CPn_0303 342982 341921
CT303 hypothetical protein
REOKGLHHMDVSRKINRHTOFFVDSIDGVIKNFDHKPSEDKSRDHEELEEKLLTITKRIV
ASAQEFQNRKTDSKNYYLKKTOWLPFKNEELEQTKELFAMLTSMDKKIAQLFFYSPGCSS
DWVEFTEVICHLNDSIGIGGVLLCCGLFEQQCEHVVTVNKKLDLPLLLGTTVVNSLRYYL
TYRNISLLNCOSMSELGKELGDVLKQHGVAFTLIFKEIVDIDLLNYVKLIQGLKRSGNIQ
ARIYDNDVPTLPSVSSSFIALRYSLANTIRGLALHVDFSSLKFISPSILSNTEHTAKALN
SGGECFIFSNLDEFNLGMKIVMQLLRTGKISPEILNKNIMKILMIKRRVRSLYI

34 3091 CPE_0304 344091 344159
pdhA/odpA-Pyruvare Debydrogenase Alpha
DQRELPRRLPYKKVMDSSAPYNIASO/TEKSTVERTILDLYGPASCIKFLKQMVLIREFEA
RGERAYLEGLVGGFYHSYAVQEAVATAAIANTGLDPWVFSSYRCHALAILLNIPLQEIAA
ELLSKETGCALSROGSHIMG-GPNFFGGGILVSKOITELAAGAAFTIKYQFQKNRVSLCFIG
UGAVAG/VHIETLINFYLSHOLFLML IENN WOMGTSLINRAVAKQPIAESQGSSYDIRAV
TVNGFDLFNSLLSFREAYRYMVDTEGFVLVEGLOSKFRGHGILDDNH/YGKREMOCLFKK
DPIVLAPIWLIRLEVLTEEEFONIRQEGKTAVLEAFSNAKLGSDPSYTTLEEGVYA 344158

344142 345137 CPO_0.00

J44142 JOHES

PhBP-odp8-TyTHVATE Debydroarenine Beta

RKEOMPRIKTHETREALREATDERMODENNY-THOREVORYROAYKVTKOLLOKWIPKRV

LDAPTSEAAFSGTGTGAALSGLPPTTEFFMSWIPSFVALDQTTSHAAKMIPMTXSKFSVPT

VFRSTRKAAAQVSCQHSHCVESTVANTFSLTTTAFSSHYYDAKGLASSATRNNNYVLFLEN

ELEYNLKGEVPTEEYLVPTGKAHRVQFSHDLTTTYSHMVSTTKEACHLAKKRWILSTET

345136 OPT_USUA pdnC-olhydrolipoamide Acetyltransferase GKFVIGLLKMPKLSPTMEVGTTVKWHKKSNDOVSFODVIVEISTDKAILEHTANEDGWIR EILRHEGEKIVIGTPIAVISTEANEPFNLEELLPKTEPSNLEASPKGSSEEVSPATTPOA ASATFTAVTFKPEPPLGSPLVFKHVGTTNNLSPLARQLAKEKNIDVSSIQGSGPGGRIVK KERRYALPPOLARE MERCHEMPROMISEEN DE EFFOLANDIMAK DE EFFOTROOVY AMERIKAKELAN DEKLAMET PRAMALALMEFOTEN OMEROMAK TYRFOT DELIA VALPOGITTPITROADRKNICMISAETKSLALKARNOSIQOTEYKOGSFOVSNIGMTGIT ETTATIVIPODATILAVGSVTEQALVLDGEITIGSTCNLTLSVDHRVIDGYPAAMFMKRLQ KILEAPAVLLLN

348998 346515

GPT_USUY
glgP-Glycogen Phosphorylase
NGCIVEDFSSFDKNKVSVDSMKRAILDRLYLSVVQSPESASPRDIFTAVAKTVMEWLAKG
WLKTQNGYYKNDVKRVYYLSMEFILGRSLKSNLLNLGILDLVRKALKTLNYDFDHLVEME
SDAGLGNGGLGRLAACYLDSMATLAVPAYGYGIRYDYGIFDQRIVNGYQEEAPDEWLRYG
NPWEICRGEYLYEVRRYGRVIHYTDSRCKQVADLVDTOEVLAMAYDIPIRGYGNDTVNSL
RLWQAQSPRGFEFSYFNHGNYIQAIEDIALIENISRVLYPNDSITEGQELRLKQEYFLVS ATIQDI IRRYTKTHICLDNLADKVVVQLNDTHPALGIAEMHILVDREELPMOKAWEMT ATIQDI IRRYTKTHICLDNLADKVVVQLNDTHPALGIAEMHILVDREELPMOKAWEMT VIENYTNHTILPSALERWPLDLFSKLLPRHLEIIYEINSRWLEKVGSKYPKNDDKRSLS IVEEGYGKRINMANLAVVGSAKVNGVSSFHSQLIKDTLFKEFYEFFPEKFINVTNGVTPR RWIALCNPRLSKLLNETIGDRYIIDLSHLSLIRSFAEDSGFRDHWKGVKLKNKQDLTSRI YNEVGEIVDPNSLFDCHIKRIHEYKRQLMNILRVIYVVNDLKENPNQDVVPTTVIFSGKA APGYVMAKLIIKLINSVADVVNODSRVNDKLKVLFLPNYRVSMAEHIIPGTDLSEQISTA GMEASGTGNMKFALNGALTIGTMDGANIEMAEHIGKENMFIFGLLEEGIVOLRREYCPOT ICDKNPKIRQVLDLLEQGFFNSNDKDLFKPIVHRLLHEGDPFFVLADLESYIAAHENVNK LFKEPDSWTKISIYNTAGMGFFSSDRAIQDYARDIWHVPTKSCSGEGN

CPn_0308 349213 349596
No robust homolog present in Genebank/EMBL as of 11/7/98
FFTOENMATVACTPOTTOPOPSVSHKATHRYCSWYFFKPILVSLGLLLASLTTLGLVIA
SGVTLSLGIGIVLAIQIVLAGIALVLAFNHIRQFKQARTAELNSMKMISAPAAATVOKQK

CPn_0309 350977 349595
CT309 hypothetical protein
FMRAWEFFLLLOEKEIGTNTYDKWLRSLKVLCFDACNLYLEAQDSFQITWFEEHIRHKVK
SGLYNNNNKPIRVHVTSVDKAAPFYKEKOMQOEKTAYFTMHYGSVNPENTFSRFLVTPEN
DEBFRVLQEFTKSPDENGGYTFNPIYLFGPESGKTHLMQSAISVLRESGGKILYVSSDL
FTEHLVSAIRSGEMOKFRSFYRNIDALFIEDIEVFSGKSATQEEFFHTFNSLHSEGKLIV
VSSSYAPVDLVAVEDRLISRFEWGVAIPIHPLVQBELRSFLHRQVERLSIRIGETALDFL
IYALSSNVKTLLHALNLLAKRVMYKKLSHOLLYEDDVKTLLKDVLEAAGSVRLTPLKIR
NVÄGTYYGVSQESILGRSQSREYVLPRQVAMYFCRQKLSLSYVRIGDVFSRDHSTVISSIR
LEDGKIEENSHDIHMAIQDISKNLNSLHKSLEFFPSEEMII

CHEO310 353472 351049
601M-60kDa Inner Membrane Protein
YFPDELSLIFRVYQMXRTLLFYSLIGIAFVGCQIFFGYNEFRSCKNLAEKQRKISEQTLA
AVESVGLSVASWDTDVNGEHKNNYAVROGDKLFLLHNGEAAQSYYSSGESWSFVDHKCG
FDHEHLALYRQQSSFNPTNTGKVFLPTNHEGLPVLVVEFRNNKEPLVFLGEYAQGRISN
KDSHIFGTALVFWRGSDYIPLGLVDSREERLVSLDLPITRAVIFGNDODSAKSSDTANH
YVLHNDYMOIIVSEESGSIEGINLFFASTNNKSIVNEIGFDRDLASEKSPELAFPGLSSK
LPPGQQAKNSIGGYYPLLRRGLLSDSKKLLPLEYHALAVVSGRELAFPVALRYRVLSYTP
HSTOLFSL DESVOKUVEL DEPUBEFBUNGEN LPDGQQAKNSIGGYYPLLRRGLLSDSKKLLPLEYHALNVVSGRELAFPVALRYRVLSYTP HSIQLESLDRSVGKYVKLPENPEREKPYVFETAITLTKETEDVWTSGYPEVEIMSNASAP TIGERVIKKNKGSLDKVKLPKVKEPLAIRRGYYPGWILNSNGYFGIILTELSEIASGYS LYISGSTAPTELSAISPKNQLYPVSKYPGYETLLPLPKDAGTHRFLVYAGPLAEPTLKVLDKBTAFTAFFALFIIMFFKLVTGSWGISIILLTVLKLLLYPLNAWSIRSMRRMQILSPYIQQIQQKYKNEPKRAQMEIMGLYKTNKVNPITGCLPLLIGLPFLIAMFDLLKSSFLLRGASFIPGWIDNLTAPDVLFSWQTSIWFIGNEFHLLPILLGIYMFLOGVKYTSLKKKGYYTQQGXQOYMGNMAILFTAMFYNFPSGLNIYWLSSMILGVGGWITNKILDSKHLKNEVVLNNKKHR

354453 353575 CPHEUSII 354453 35575
CTJH hypothetical protein
DMAMMAVIYWDRSKIVWSFEPWSLRLTWYGVFFTVGIFLACLSARYLALSYYGLKDHL
FSKSQLRVALENFFIYSILFIVPGARLAYVIFYGWSFYLQHPEEIIGIWHGGLSHGGV
GFLLWAAIFSWIYKKKISKLTFLFLTDLCGSVFGIAAFFIRLGNFWNGEIVGTPTSLPW
VVFSDPMGGVGGVPVHPVQLYEGISYLVVSGILYFLSYKRYLHLGKGYVTSIACISVF RFFAEYVKSHQGKVLAEDCLLTIGQILSIPLFLFGVALLIICSLKARRHRSHI

CPn_0312 354518 354976
CT101 hypothetical protein
CTMARNIKYFLILFPGILWISAGMKLLKATAIALDPLSSFFTYCLLSMVSWGLASLKHR
YLLSKTIRKQLSLSSEFFSQKTWIAYIKQTFISRRFLIMVIMIAFSLVLRRYISNPQAL
FVIRATVGYALIKTAIAYFSKLQNALMENPEGN

acpS-Acyl-carrier Protein Synthase WKILKEISANSMEIIHIGTDIIEISRIREAIATHGNRLLNRIFTEAECKYCUEKTDPIPS FACRFAGKEAVAKALGTGIGSVVAWKDIEVFKVSHGPEVLLPSHVYAKIGIKKVILSISH CKEYATATAIALA

CPn_0314 356285 355353

trxB-Thioredoxin Reductase
MIHSRLITIGSGPSCYTAAIYASRALLHPLLFEGFFSGISCOQLMTTTEVENFPGFPEGI
LGPKLMNNMKEQAVRFGTKTLAQDIISVDFSVRPFILKSKEETYSCDACIIATGASAKRL
EIGAGNDEFWOKCYTACAVCDGASPIFKNKOLYVIGGDSALEEALMITRYGSHVYVWH
RRDKLKASKAMEARAQNNEKITFLWNSEIVKISGDSIVRSVDIKNVTQEITTREAAGVF
FAIGHKFNTDFLXQUTLDESGYIVTEKGTSKTSVPGVFAAGDVQVKYYRQAVTSAGSCC
IAALDAERFEG

356977 358716 il Ritensomal Protein

1

KYPPOKRVICK IVKLLPYC EGIEGLIH ISEMSWVKN IVDPSEVVNKCDEVEAL I LSIOKDECK ISLGLKOTER EEKYP ICLHVNAE IKNLTNYGAFVELEPGIEGLIH ISOMSWIKKVSHPSELFKKONSVEAV ILGVOKESKK ITLGVKQLSSNPWNE I EAMFPAGT VISCOVIKITAFGGAFVELQNOI EBLI HVSELSOKPFAK IEDI ISIGENVSAKVIKLDPDH KKVSLSVKEYLADNAYDQDSRTELDFKDSQGPKERKKKGK

CPN_0316 358784 360121
nusA-N Utilization Protein A
LETERLIVYP.IMBERHIAAIJOMEKERGIORETIIGAIEGALKIAAEVTLPODANIGEN
LETERLIVYP.IMBERHIAAIJOMEKERGIORETIIGAIEGALKIAAEVTLPODANIGEN
LEGERIAFIEVE EMERICAL JAHOUKERTELBE AREEDBIO 3000YMPOVEPUDDEGERAFE
ARKOITOKUKHAAEKUVIYJEVENIKWINETLEGOVVNEFAKGSNLIIOLGKYEAILPTRAYE
KTEKHKIGDKIYALLYEVQESENGGAEVILSRSHAEFVKQLFIQEVPELEEGSVEIVKIA
REAGYRTKLAVRSSDPKTUPVGAFVGHEGSRVKNIIRELNDEKIDIVNYSPVSTELLQNIL
LYPIEIGKIAILEDDKVJAIVVANGVANGVANGVANGVANGVANGVANGVANGVANGVENSTYNKL
LEIQRLQLAEFDSPHLDOPLEMEGISKLVIQNLEHAGYDTIRRVLLASANDLASVPGISL
ELAYKILEQVSKYGESKYDEKPEIED

360045 362750

CPn_0317 360045 362750
infB-Initiation Factor-2
SLLIRSLSKSANMERVKLTKNIKLKIKNAQLTKAAGLDKLKQKLAQAGSSEAKSSSEKPS
AKEKSVKVALAATSPTASAĐQASPESTSRRIRAKNRSSFSSSEEESSAHIPVDTSEPAP
VSIADPEPELEVVDEVCDESPEVHPVAEVLPEQPVLPETPPQEKELEPKPVKPAEPKSVV
MIKSKFGPTGKHIMHLLAKTFKAPAKEEKVVAGSKSTKPVASDKTGKPGTSEGGEQNNRE
KQFNPANRSPASPKRDAGKINLTDFRDRSKKSDESLKAFTGRDRYGLHEGGEDRWRKK
RVYKPKKHYDEZSIQRPTHIKISLPITVADLAAEMKLKASEVIQKLFIHGMTYVVNDILD
SETAVOFIGLEFGCTIDIDYSEQDKLCLSNDTVRDEIQSTDPSKLVIRSPIVAFMGHVDH
GKTTLIDSLRISNVAATEAGAITOHMGAFCCSTPVGDITILDTPCHEAFSAMRARGAEVC
DIVVLVVAGDEGIKEOTLEATEHAKAADIAIVVAINKCDKPNFNSSTIVRQLSEINLLPE
AWGGSTVTVWTSAKTGEGLSELLEMLALQAEVLELKADPSARARGLVIESELHKGLGPVA
TVLIONGSIKLGEALVFNDCYGKVKTMINENNELMKEAGPSIPVLITGLSDIPKAGDPFF
VVKNEKTARDIIEARSAGOQRFALQOKKPPNFDSMLONKKTLKLHIKADVOGSIEALVSS
ISKIKSEKVDVEILTNSVGEISESDIRLAAASKAVLIGFHTGIESHAEPLIKSLGVVVEL
FTVIYHALDAIKEIMTSLLDPIAEEKDEGSAEIKEIFRSGVGSTYGCIVTEGIMTRNHK FTVIYHA(DAIKEIMTSLLDPIAEEKDEGSAEIKEIFRSSOVGSIYGCIVTEGIMTRNHK VRVLRNYEILWKGTLSSLKRVKEDVKEVRKGLECGILLEGYQQAQIGDVLQCYEVIYHPQ

CPn_Q118 362704 363126
rbfa/Ribosome Binding Factor A
vmsynwmxusithknynukrychtenrikrvnalloeaiakvilkdvkhpkisnlwitvt
rvs_skolhsarvyvsymphentkeealealkvsagfiahrasknvvlkyfpelhfyldd
ifspodyienluwqioekeks

CT_0319 363133 363879
truB-tRNA Pseudouridine Synthase
TIFFGMINTIKDMINDLAVELKEGILLVDKPOGRTSFSLIRALTKLIGVKKIGHAGTLDP
FATGVMVMLIGRKFTRLSDILLFEDKEYEAIAHLGTTTDSYDCDGKVVGRSKKIPSLEEV
LSAAEYFGGEIQQLPPMFSAKKVQKKLYEYARKGLSIERHHSTVQVHLQITKYEYPLLH
FVVSCSKGTYIRSIAHELGTMLGCGAYLEQLRRLRSGRFSIDECIDGNLLDHPDFDISPY

CPn_0320 363824 364783

LRDAHGNSI.

TiDF-FAD Synthase
TTPISIFLPTYEMPMEIAYSLTSSFSVDSVTVGFFDGCHLGHSNLLSILTSYSGSSGVIT
TEDSHPQTVLSLNHYKLINTKEERLQLLQTFPIDWLGVLTFDLNFANQSAEEFLTLLHRNL
KCKRLILGYDSCIGKEQOSNTEALDTIGKPLGIEVIKIPPYRMDNIVVSSKAIRQFLSAG NLECAHRFLGHPYAISGKITEGSGIGGSLGFATINLPREESLIPLGVYACEIRYDSTTCQ GVMNLGTAPTFGRESLYAEAHIFSFAENLYGKEVSIIPRKFLREEKKFOSKETLIRAIEK DILDAODWEAKGSENVEGTA

365900 364767

ychF-GTP Binding Protein YSKKHVIIFIFRCLMSHTECGIVGLPNVGKSGLFNALTGAQVASCNYPFCTIDPNVGIVP VIDERLEALAKISNOKIIYADMKEVDIAGLVKKASDGAGIGNFLISHIRETHAIAHVVP
VIDERLEALAKISNOKIIYADMKEVDIAGLVKKASDGAGIGNFLISHIRETHAIAHVVP
CFDDPDVTHVSGKVNPVEDIEVINLELIFSDFSSAKNIHSKLEKLAKGKREVGALLPLFD
TIIAHLEKGLPLRTLELTPEQIVALKPYPFLTMKPMFYIANVDESSLPDMDNDVVAAVRE
VAAKENSKVVPICVRIEEEIVSLPIEERLEFLMSLGLEKSGLRHLVRAAVDTLGLISFT
TGPOGSRAWTVVRGSSAWEAAGEIHTDIQKGFIRAEVITFEDMIECQGRAAARELGKLHI EGRDYIVODGDTMLFLHN

CPn_0322 366231 367328
yscU-YopS Translocation Protein U
SNLGNSMGEKTEKATPKRLRDARKKGQVAKSQDFPSAVTFIVSMFTAFSLSTFFFKHLGG
FLVSMLSQAPTRHOPVITLFYLKNCLMLILTASLPLLGAVAVVGVIVGFLIVGFTFSTEV
FKPDIKKFNPIENIKQKFKIKTLIELIKSILKIFGAALILYITLKSKVSLIIETAGVSPI ITAQIFKEIFYKAVTSIGIFFLIVAILDLVYQRKNFAKELKNEKFEVKQEFKDTEGNPEI KGRRRQIAQEIAYEDSSSQVKHASTVVSNPKDIAVAIGYMPEKYKAPWIIAMGINLRAKR ILDEAEKYGIPIMRNVPLAHQLLDEGKELKFIPESTYEAIGEILLYITSLNAQNPNNKNT NOPDHL

367322 369460

CPn_0323 367322 369460
lcrD- Low Calcium Response D
SFIRMKLINFYSRTLCGOTALIMINKSSDLILALMMMGVVLMIIIPLPPPIVDLMITINL
SISVFLLMVALYIPSALQLSVFPSLLLITTMFRLGINISSSRQILLKAYAGHVIQAFGDF
WCGNYVVGFIIFLIITIIQFIVVTKGAERVAEVAARFRLDAMPGKQMAIDADLRAGMID
ATQARDKRAQIQKESELYGAMDGAMKFIKGDVIAGIVISLINIVCGLITOVAMHGMLAGA
AAHVYTLLSIGDGLVSQIPSLLIALTAGIVTTRVSSDKNTNLGKEISTQLVKEPRALLLA
GAATLGVGFFKGFPLWFSILLALIFVALGILLITKKSAGKKGGGSGATTVGAADGAA
TVGONPDDYSLTLPVILELGKDLSKLIGHKTKSGQSFVDDMIPKMRQALYQDIGIRYPGI
HVRTDSPSLEGYDYMILLNEVPYVRGKIPHHVLTNEVEDNLSRYNLPFITYKNAAGLPS
AWVSEDAKAILEKAAIKWYTPLEVILLLLSVFFHKSGQEFLGIGEVRSMIEPFMERSPPDL
VKEVTRLIPLQKLTEIFKRLVQEDISIKDLRTILESLSEWAQTEKDTVLLTETVRSSKLL
YLJEKFSQQSAISVYLLDTEIEBMIRGAIKGTUSYSKLALDPDSVNLILKSMRNTITPT
PACKYOPPVLLTAIDVRRYVRKLIETEFPDIAVIGYQEILPEIRIQPLGRIQIF

Circ 9:24 300468 370688

CTH_20_224 H0468 370688
CTH_20_224 H0468 370688
CTH_20_224 H0468 H0264 H

AEGIOTO:DUNFVKVAEGYHKIINDKFPTAGKVEREVI QTGGRUFGGADKRQQUJAMIANALDAVNINNEDYPKAS

VDSVTGVLNLFFSALR

CPn_0325 370536 371148 CT325 hypothetical protein KRIAMONOYEQLLESLAPLLNTTLAPDKNNSCLIRFSDTHVPVQIEEDGNSGDLAVSTLL GTLPENVFRERIFKAALSVNGSFQSSIKGILGYGEVTQQLYLSDILSMNYLNGEKLFEYL KLFSLHAKIWMESLRTGNLPDLHYLGIYYVA

CPn_0327 372927 373211 r128-L28 Ribosomal Protein RIHRKUMSRKCPLTGKRPRRGYSYTLRGIAKKKKGIGLKVTGKTKRRFFPNMLTKRLWST EENRFLKLKISASALRHIDKLGLEKVLERAKSKNF

CPn_0328 373220 374992

CT085 hypothetical protein

LKYREIFMSFLRRHISLFRSQKQLIDVFAPVSPNLELAEIHRRVIEDQGPALLFHNVIGS

SFPVLTNLFGTKHRVDQLFSQAPDNLIARVAHLISSTPKLSSLWKSRDLLKRISSLGLKK

AFREFFPVSMSSVNLDHLPLLTSWPEDGGAFLTLPLVYTESPTLTTPNLGMYRVQRFNQ

NTMGLHFOIOKGGGMHLYEAEOKKONLPVSVFLSGNPFLTLSAIAPLEBNVSELLFATFL

GGAKLLYKKTNDHPHPLLYDAEFILVGESPACKRAPEDFFGDHFGYYSLQMFPEFFKCHK

JYHRKDAIYPATVVGKPYQEDFYIGNKLQEYLSPLFPLVMFQVRLKSYGESGFHALTAA

VVKERYWRESLTTALRILGEGQLSIKFLMVIDQEVPLDRFSVVLETILERLQPPDLLIT

FSETANDTLDYTGFSLNKGSKGIFMGIGKAIRDLPHGYQGGKIHGVQDIAFFCRGCLVLE

TSLEDRCIKSLLHHPDLKSWPLIILADNLRETIQSEKDFLWRTFTRCAPANDLHALHSHF

ATHRPNYNFPFVIDALMKPSYPKEVEVDPSTKQKVSERWHAYFPNKETFYI

CPE-0329 375085 376146
Phopholipase D Superfamily (leader (33) peptide)
Phopholipase D Superfamily (leader (33) peptide)
MMRKROKDKLKICVIISTLILVGIFARAPRGDTFKTFLKSEEAIIYSNQCNEDMRKILCD
ALEHADEE IFLTYNLSEPKIQOSLFTROAQAKNKVTIYYQKKFLIQILKQASNVTLVEQP
PAGRKLMHOKALSIDKKDAWLGSANYTNLSLRLDNNLILGHHSSELCDLIITNTSGDFSI
KDOTGKYFVLPQDRKTAIQAVLEKIQTAQKTIQVAMFALTHSEIIQALHQAKQRGIHVDI
ITBRSHSKLTFKQLRQLNIKNDFVSINTAPCTLHKKFAVIDNKYLLAGSINWSKGRFSLN
DESSIILENLTKQQNQKLRMIWKDLAKHSEHPTVDDEEKEIIEKSLPVEEQEAA

0330 376930 376202

CÉN_0330 376930 376202

CÉN_0330 prothetical protein

FISIEMILLISROLFSVLPSRFODLHVYRFKESLKLLOFMIMVGGEIVVVLAEIKEEDLR

RKÉBVRKRREKNYLRIFRVLSRFDVMRIIRFDPYGALSAQSIAKDSRONSPLVEKISEEN

ATNEAIRIALLAIGDREQEEKKQRHRYKLLGQKQAKVLLSQLRHVHLDFKKLYCDSKKKA

DÖEKDEKNKQKRSIKVTKKKKGISLGAAASQAIAAAAEAWVIARNKGVLETASTLFYQKD

376701 378452

CÉRLO331 378452 376701
CT082 hypothetical protein
IGRI HMYSGGGGVOPSSPGKMNPALGGEQAEGPSPLKESIFSETKQASSAAKQESLVR
SGSTGMYATESOINKAKYRKAQDRSSTSPKSKLKGTFSKMRASVQGFMSGFGSRASKVSA
KRABDSGEGTSLLPTEMDVALKKGNISPEMGGFIDASGMGSSSSISQLSLEALKSSA
FSGARSLSLSSSESSSVASFGSFQKAIEPMSEEKVNAMTVARLGGEMYSSLLDPMYETSS
LVHERMATCHEGMIDLSDLGGEVSTAMTSPRAVEGKVKVSSDSPEANPPGIPMSTLS
PTÜRSGDKSKHKSFGIEKSTNHTNFSPLREGTVKSAEVKSLPHESYMFPKPSTLSPTBSGKSKKHSFGIEKSTNHTNFSPLREGTVKSAEVKSLPHESYMFPKPSTVSREE
PEAVVKESTAFKMPENSSONFLPIAVESVFFKESGTGGALGSDAVSSYHFLAQRGVSLL
APLPRATDDYKEKLEAHKGPGGPPDPLIYQYRNVAVEPPIVLRSPQPFSGSRLSVQGKP
EAASVHDDGGGGNSGGFSGDQRRGSSGQKASRQEKKGKKLSTDI

378676 378536 CPn_0332 CHLTR T2 Protein

YLDSRIRVIPLARORCTLLHLLAVLCPPISFFTQGVSPCVFFCFLD

378800 379117 CPn _0333

TELB VDEFVFFMGKPKKSRTDRALAGEIQKKSTEVLKKPARIKAKORRKFLIAKEOKTLKHR AGEYDQLVRSLLDSQKKDTDKVLIFNYENGFVFTDKDHFSKY9IRL

379308 379823

CPn_0334 179308 379823
CTOT9 similarity
TMSVHITPRKCFILCILSMFTLPTLFPKAHLILFSPYIVVCFYCFSKDKGLVLALGCGVL
SDLALGSGGVFLLLYPLTALITHKAHLIFSKESKAALVYVMIFYGVFLLLTIPMCALFG
HEVRWSIDVLMIPLKCSFLDNLIFTSVIYILPCAINSGTHKMISFFRRLVCY

CPn_0335 379808 380674

toID-Methylene Tetrahydrofolate Dehydrogenase
EIGMLLPGIPAAEKILDRIKKETISOSPTSPOLAVVZIONDPASEVYVOMKVKKATEIGII
SKANKLPGDSTLSSVLKLIERLNODPSIHGILVOZPLPKHLDSEVILDAISPOKDVOLH
PVNMCKLLLCNFDGLLPCTPAGIIELLNYEIPLRGPHAAIVGRSNIVGKPLAALMMOKH
PVTNCTVTVLH:GSENLPBILKTADIIIAALGYPLFIKETMVAPHAVIVDVGTTRVPADN
AKGYTLL/JDVDFNNVVTKCAAITPVPGGVGPMYVAMLMSNTWRCYQNFS 379808 380674

380569

yojl rodkmodhedownewshicryaheoymanlekffluulolokooooktitiedeomit fyhtuuttilokakekasloooolokuthkootopeksuutiikkideoretuuttilokakekasloooolokuthkootopeksuutiikkootopeksuutiikkideoretuutekstuulohekootopeksuudiikkootopeksuudiikkootopeksuutiikkootopeksuutiikkootopeksuutiikkootopeksuutiikkootopeksuutiikkootopeksuutiikkootopeksuutiikkootopeksiikusothieoretusootopeksuutiikkootopeksiikuutiikkootopeksiikusotopeksiikusothieoretusootopeksiikisootopeksiikusotopeksiikusotopeksiikisootopeksiilistootopeksiikisootopeksiilistootopeksii

CPD_033.
smpl Small Protein
smpl Small Protein
LEEIFFGNGKRILLIVERKNCFLLYWFLSPIMGEDLMAGKEIVSNPKALRNYEVIETL
LEAGIVLTGTEIKSLENDHOGNLGDAYVIVSKGEGWLLNASIAPYRFGNIYNHEERRKRKLL
LHRYELRKLEGKIAGKGWTLIPLGMWLSRGYVKVRLGCCRGKKAYDKRRTIIEREKEREV

383375 ·in

CTING INVOCEMENT OF ANY ENMERCEMENT OF ANY AKVYEKGALSIPOKPER PRICE PONTE LEVILIBRI LETYIDELVETATEUTVOTROVTK ENMERGURINE FOLL EFFOLTE LETTER ENTER LETTER ENTER LETTER ENTER AKVYEKGA IS I POKRE FOLT KEUTEANLE I SSUADEMO I TUGSSOCRELLSMEKED FREI POLT AKVYEKGA IS I POKRE FOLT KEUTE EN STATE EN STATE TUGT KELAK POT GERALAK I DAEVILDKEFSGEVI I PI KAVEET I KNCSDEGEAA I FLOODKI AVECDNTLLITKLLSG EFFDFSPVISTESNVKLDI KREELITLIK OVALFTNESSHSVKFSFLØGELTLTANCTKV GECKVSMAVNYSGELLE JAFNEFFFLDILKHSKDELVSLGISDSYNPGI I TDSASGLFVI MANDELHOL

383405 384034

CPn_0339 383405 384034
CT339 hypothetical protein
vTTLPMFMKICSLKLMNFRNHSDLEISLAPKLNYAQGKTNLLEALYVLSLGRSFRTOHLT
vTTLPMFMKICSLKLMNFRNHSDLEISLAPKLNYAQGKTNLLEALYVLSLGRSFRTOHLT
vTTLFMFMKICSLKLMNFRNHSDLEISLAPKLNYAQGKNICHOLPIKTLSQLIGKVPIVLFSS
KDRLLISGAPADERLFLNLLLSQCDNHYTLCLSYYHRALQQRNALLKSKQTSTVASGMS
WSNTAPTYPSNGFSVVRNFQIYPKNFGLTT

383842 384156 CPn 0340

(frame-shift with 0339)
PLYPLLIVISSRSSAEKCSLKKQANLNRGLWDEQLVKHGTYLSIQRFLCSQKLSDLSKEL
WSNNLKEQLALKFKSSLIKNSDISETAVAEEFHKQLSISLPRDLE

41 384160 384495 Shift with 0340)

(frame-shift with 0340)
GSTSVEPHREDFLLTMNOMPVSQFSSEGQKHSLLAILRLAECLYLKQSHHVSPLVCLDDI
HAGLMARVGQLLDPAPTLGQTLITSTHMHGELPKTSLVLSIENAQVSEQII

CPT_0342 384619 385062
predicted OMP [leader (19) peptide]
HMKKPLLTILFLAVCNPLFSETSVIOTLPSGIGGLKETSKOKESVVCVHAFLRSYTSLKP
FARVLEKEHYDVF IMNYETRKFTLEKHAEHLNRLLKKIAELKPGVPINFVTHSIGGVIVR VALAHPDCPEEAKKEKLFSWLLRTQGLH

385595 384999

CPn_0343 384999 385595
(frame-shift with 0342?)
LPRRSQKRKAILMAPPNAGSTLARRYRCVKFVQFVFGGKLGRQLLTYCPTKMLNVGKLPS
SLDVLLLSGNRHSKFLPFRLPYENDGKVCTIETKLDTPHKAYVIHTSHTYIITNKSLYL
MKEFLKEGNTTPIIEHVPEAALEQTVMEDKQKNSRLKPYPNQDIYVIHCFGSRPYNLYGF PKKWSLNQKNEINPEKLEK

CPn_0344 387432 38558

yael-Metalloprotease
SSRWTIIYFILAALALGILV/LIHELGHLVVAKAVGMAVESFSIGFGPALFKKRIGGIEY
RIGCIPFGGYVRIRGMERTKEKGEKGKIDSVYDIPOGFFSKSPWKRILVLVAGPLANILL
AVLAFSILYMNGGRSKNYSDCSKVVGWVHPVLQAEGLLPGDEILTCNGKPVVGDKDMLTT
SLLEGHLNLEIKRRGYLTVPSKEFAIDVEFDPTKFGVPCSGASYLLYSNQVPLTKNSPME
NSELRPNDRFVWMDGTLLFSMAQISGILNESYAFVAVARNDKIFFSRQPRVLASVLHYTN
YLRNELIDTQYEAGLKGKWSSLYTLPYVINSYGYIEGELTAIDPESPLPOPOGENLOLGDR
ILAIDGTPVSGSVDILRLDPVOPRPWIDVYSGSESLDKQLEVAKKINKDKDRYYLERLD
LINHLGESHPVEVAGPYRLLDPVOPRPWIDVYSSESLDKQLEVAKKINKDKDRYYLERLD
AEKQKPSLGISLKDLKVRYNPSFVVMLSNITKESLITLKALVTGHLSPQWLSGPVGIVQV
LHTUMSVGFSEVLFWIGLISMNLAVINLLPIPVLDGGYILLCLWEIVKRRRLNMKIVERI
LVPFTFELLIFFIFLTFODLFRFFG LVPFTFLLIIFFIFLTFQDLFRFFG

388587 387436

CPn_0345 388587 387436
CT345 hypothetical protein
LKVACLKHLAVLGSTGSIGROTLEIVRRYPSEFKIISMASYGNNLRLFFQQLEEFAPLAA
AVYNEEVYNEACORFPHMOFFLQGEGLYQLCIMDTVTTVVAASSGIEALPAILESMKKGK
ALALANKEILVCAGELVSKTAKENSIKVLPIDSEKNALYOCLEGRTIEGIKKLILTASGG
PLINKSLEELSCVTKGDVLNHFIMMMGSKVTVDSSTLVNKGLEIIEAYWLFGLENVEILA
VIHPOSLIHGMVEFLOGSVISIMNPPDMLFPIQYALTAPERFASPROGMDFSKKQTLEFF
PVDEERFPSIRLAQQVLEKQGSSGSFFNAANEVLVRRFLCEEISWCDILRKLTTLMECHK
VYACHSLEDILEVVGEARALAQEI

CPn_0346 389690 388704
070-troD/ytgD-Integral Membrane Protein
KKGSIMALGPSPYYGVSFFQFFSVFFSRLFSGSLFTGSLYIDDIQIIVFLAISCSGAFAG
TFLVLRKMAMYANAVSHTVLFGLVCVCLFTHQLTTLSLGTLTLAAMATAMLTGFLIYFIR
NTFKVSEESSTALVFSLLFSLSVLLVFMYKNAHIGTELVLGNADSLTKEDIFFVTIVIL
ANAVITIFAFRSLVCSSFFSVFASSLGIPIRLVDYLIIFQLSACLVGAFKAVGLMALAF
LIIPSLIARVIAKSIRSLMAWSLVFSIGTAFLAPASSRAILSAYDLGLSTSGISVVFLTM
MYIVVKFISYFRGYFSKNFEKISEKSSQY

391078 389678

CPn_0347 391078 389678
069-troC/ytgC-integral Membrane Protein
TFGTNPEALSRKTIWIVLIMLSCVPSDTIFLSSFLAVTLICMTTALWGTILLISKQPLLS
ESLSHASYPGLLVGALMACYVFSLQASIFWIVLFGCAASVFGYGIIVFLGKVCKLHKDSA
LCFVLVVFFAIGVILASYVKESSPTLYNRINAYLYGQAATLGFLEATLAAIVFCASLFAL
WMWYRQIVVTTFDKDFAVTCGLKTVLYEALSLIFISLVIVSGVRSVGIVLISAMFVAPSL
GARQLSDRLSTLLILSAFFGGISGALGSYISVAFTCRAIIGQQAVPVTLPTGPLVVICAG
LLAGLCLLFSPKSGWVIRFVRRKHFSFSKDGHLLKVFWHISHNRLENISVRDFVCSYKY
QEYFGPKPFPRWRVQILENRGYYKKEODYYRLTKKGRSEALRLVRAHRLWESYLVNSLDF
SVECVHELABELEHVLTEELDHTLTEILNDGYYDDHROITIBNKKFFV SKESVHELAEEIEHVLTEELDHTLTEILNDPCYDPHPQIIPNKKKEV

CPn_0348 391815 391027
068-troB/ytgB-ABC transporter ATPASE
ECWILLNUYDETFISOVANLCYNTEHAAVILYH DEGLGYGGLTATLGPNGAGKOTLLKASLG
LLYPSGGTY/FFNQKFKKYRQFLYYMPQRACYLWIDFPMTVLDLALMX:YSYKCMWGRLGS
DDRREAFHILERYGLESVAPRQT/GOLXXIQQY/AFLAKALMQKADLYLMDELFIATDMAS
FYTSVGVLQELRQQGKTIVVVHHDLSHVRQLFDHVVLLNKRLTCGGPTDECLNIDTTFQT
YGGETELLEOTLKLGRGKGFGGY.

301700 cpn_0349 - 孫紀紀 391790 - 067-troA/yrgA-Solute Protein Binding Family WILKMAS:REMDAKMGYIFKYMRWIFCFVACGITFGCTN
VERVYGNRLATAVLIKGSLDPHAYEAVKGDKDKIAGSA
PNGYKLGERLIARGAFVPLEEDGICDPHIWMDLSIWKER
NSEELVCEWSILDSWAKOCLSTIPENLRYLVSGHNAFSYFTRRYLATPEEVASGAFKA
ISPEGLSPEAGISJROIMAVVDYINEHDVSVVFPEDTLNQDALKKIVSSLKKSHLVRLAQ NSRPCILSMNRMIHDC KPLYSDNVDDNYFSTFKHNVCLITEELGGVALECQR

393684 393169

CPP_0350
TOOL DESCRIBE SET OF PROFITS
FOR BRANCHERS FROM THE PROFITS SET OF STREET OF SET OF

393861 395432

CPn_0351 193861 395432
adt-adp/Atp Translocase
kikvforvnntkteekpfcklrsflwpihthelkkvlpmflmffcitfnytvlrdtkutl
kikvforvnntkteekpfcklrsflwpihthelkkvlpmflmffcitfnytvlrdtkutl
kvargsgabaipfikfwlvvpcaiiffmliyaklsnilskoalfyavctopfliffalfpt
viyplrdvlhptefadrloailppgllglvailrnwffaapfvlaelmsvmlslmfmaf
aneitkiheakrfyalfgiganisllasgraivmsklrasvsecvdpwgislrllmamt
ivsglvlmasywminknvltdprpfynpeemokokkgakpkmmkdsffylarspyillla
llviargicinlievtwksolkloypnmndysefmgnsffmgvvsvlihlfvcgnvirk
fgwltgalvtpvwylltgivffalviffnoasglvamfgttplmlavvvgaionilskst
kyalfdstkemayipldgeokvkgkaaidvvaarfgksggalioggllvicgsigamtpy
lavillfiiaiwlvsatklnklflagsalkeoevagedsapass

CPn_0352 395478 396830
No robust homolog present in Genebank/EMBL as of 11/7/98
WWGIFFINSHTNSYAFFNQKVIITVRHSGCTMKCSPLTLVPHIFLKNDCECHRSCSLKI
RTIARLILGLVLALVSALSFVFLAAPISYAIGGTLALAAIVLLIITLVVALLAKSKVLPI
PNELOKIIYNRYYKEVFFYVKTHSLTVNELKIFINCWKSGTDLPPNLHKKAEAFGIDILK
SIDLTLFPEFEEILLQNCPLYWLSHFIDKTESVAGEIGLNKTOKVYGLLGPLAFHKGYTI
IFHSYTRFLLTLISESQYKFLYSKASKNQWDSPSVKKTCEEIFKELPHMIFRKDVQTI
GPLFLFFSHGITWEQAGNIQLINPDNWKMLCOFDKAGGHCSMATFGGFLMTETNMFDPVS
SNYEPTVNFMTWKELKVLLEKVKESPMHPASALVQKICVNTTHHQNLLKRWQFVRNTSSQ
WTSSLPQYAFHAQTYKLEKKIESSLPIRSSL

CPn_0353 396893 397135
No robust homolog present in Genebank/EMBL as of 11/7/98
LRFRNIKKSLIFIKRIRYSQSGKEQKGARPFFKKSITSSLVILLLEAIFNENFSSIIQNN
FNKNFKNKNISINRIFVKFTI

CPP-0354 397062 398507

NO PODUST homolog present in Genebank/EMBL as of 11/7/98
YKTESIKILKIKTFLLIGFLLNRYNTOIDEPRKCMSNITSFVIQNNRSCNYYFELNNST
TIHTVISAILLCGALIAFLCVAAPVSYILSGALLGLLIALIGVILGIKKITPMISSKE
OVEPOELVNRIRAHYRFVSDFVSEAKFNLKDLISFIDLLNQLHSEVGSSTNYNVSEELQ
OKEDFFEGIARLKNEVRTASLKRLESAASSRPLFPSLPKILOKVFPFFWLGEFISAGSKV
VELERVKKIGGSLEEDLSDYIKFENLPTYWLIPLDFRPTNSSILNLHTLVLARVLTRDVF
OHLKYAALNGEWNLINSDLNTHKQOLFAKYHAAYQSYKHLSOPSLQEDEFYNLLLCIFKH
RYĞWKQMSLIKTVPADLWENLCCLTLDHTGRPODMEFASLIGTLYTQGLIHKESEAFLSS
LTIESSLOFKTIRRQSTNIAMFLENLATHNSTFRSLPPITVHPLKRSVFSQPEEDESSLL
IGE====

CPRE0355 39995 398591
NOTED BY A STATE OF THE PROPERTY OF THE LEPKETHHILQULTKEDVLLLKNKALQEKWDTDEVKATVERTYTTYTARGTLKTEAGGLT KETISKELLLLSLHGYSFDQLQLITQLPRDAWDWLCFVDNSTAYNLQLCALVGALSSQNL LDESSIDFDVNLGLYVIQDLKEAVQAFSASDEPKKELGKFLLRHLSSVSKRLESVLRQGL HRIALEHGNARARVYDVNFVTGARIHRKTSIFFKD

CPT=0356 400465 400109
No Ebbust homolog present in Genebank/EMBL as of 11/7/98
KQVGLFQYMNESGWDWLCDFDSQGGGFQLSRLVGLLHSSWALYEAKEGFYLPEVSLI ELTEMOLLSKPTKHGVAKDLCNVFEKHFQRFRQYLGSLDLNQRFENTFLNYPKYHLDF

401341 400469 CPN_0357 401341 400469
No robust homolog present in Genebank/EMBL as of 11/7/98
YSSHNGASMVNIQPVYRNTQVNYSQATQFSVCQPALSLIIVSVVAAVLAIVALVCSQSLL
SIELGTALVLVSLILFASAMFMIYKMRQEPKELLIPKKIMELIQEHYPSIVVDFIRDDEV
SIYEIHHLISILMKTNYFOKAPVYLQEKLLQFGELFKFDVHPSKLPNFEELLDGVELHW
LGRLVYPMVSDVTPGTYGYYWCGPLGLYENAPSLFERRSLLLLKKISFGEFALLEDGLKK NTWSSSELVQIRQNLFTRYYADKEEVDEAELNADYEQFDSLLHLIFSHKLS

401757 401578 No robust homolog present in Genebank/EMBL as of 11/7/98 EEVLSVSMKLIPTQDSIERETDSKRDKKIFTIYICSSKVLAGHFFSHLDKHNKIHESIGV

401994 403817

LepA-GTPAse
ITLOY ILKEYKIENIRNFSIIAHIDHGKSTIADRLLESTSTVEEREMREQLLDSMDLERE
RGITIKAHPVTMTYLYEGEVYQLNLIDTPGHVDFSYEVSRSLSACEGALLIVDAAQOVQA
QSLANV/LALERDLEIIPVLNKIDLPAADPVRIAQQIEDYIGLDTTNIIACSAKTGGGIP
AILKAIIDLVPPKAPAETELKALVFDSHYDPYVGIMVYVRIISGELKKGDRITFMAAKG
SSFEVLGIGAFLPKAFTIEGSLRFQQVGFFIANLKKVKDVKIGDTVFKTKHPAKTPLEGF
KEINPV/FAGIYPIDSSDFDTLKDALGRLQINDSALTIEGESSHSVGFGFRCGFLGLHL
EIIFERIIREFDLDIIATAPSVIYKVVLKNCKVLDIDNPSGYPDPAIIEHVEEPWHVNII
ITFQEYLCHIMNICALDKRGICVKTEMLDQHRLVLAYELPLHEIVJDFNDKKSVTKGYGS
FDYRLGDYPKGSIIKLEVLINEEPIDAFSCLVHRDVAESRGRSICEKLVDVIPQOLFKIP
IQAAIHKKVTARETIRALSKNVTAKCYGGDITRKRKLWEKQKKKKRMKEFGKVSIPNTA
PIEVILKID lepA-GTPase FIEVLKUD

403922 405364Ctn 0360 CTH_0.360 AD 364 AD 322 CTH_0.5 hypothetical profein VALSTHIFALGILAVMCKNLVILMILDHOGSSVCVVNRT PERTROFILKEYENHRELVGFESLE DEVNELERIHKITMLMILOAKEVODS HIALLEFILEVIJE I DOXINSYEKOSERECKELOEK GILETUYSTEXI BEIJARIESPETME XINDERAMELVAP I POŠI LAKVOGRPCCSWVOTOXIAG HYVKAVIBKI I EYGDIQLI CEAYGI LROFILKI SATAVAT LIKEWNTLELESYL I RIASEVL ALKOPEGIPVIDTILDVVOG HALYASKIISYAGGFMLDJEASKEYMGLDLGEIA ARMYPOTPLIFEMPHOPSVF HALYASKIISYAGGFMLDJEASKEYMGLDLGEIA LMWROCCIIQSAFLOVIHKGFAANPENTSLIFQEYFRGALRHAEMOWRTT/TTAIGAGLP IPCLAAAITFYDT/PTASSSMSLAGGLRDYFGAHTYERNDRPRGEFYHTDWVHTKTTERV

CPn_0361 405650 405382

EVTS-tyrosyl than synthetase
CLAIMITEDIMENT DULLETHILEBETH CONTROLLAND GEDETAPALHIGHNIOL
THER LAAD HITE CLAIM MALHENETH REPUBLICAND GEDETAPALHIGHNIOL
TUVNNADMUQELD LIPTUD KRIPKOLMUNKOTI KORVHODEGI JYTEFOYLI LOSYO
TYHLEKNYGTI LOCGSSON KNITSGI DFI RRKGLGOAYGLTYPLLTNAGKKIGKTESG
TYWLDSDLTSPELYOYLIFLPDDTI PRI ARTLITLLSNEEI QDI DRRVOTDPVAVXEFVA
QDILSA HGDIGLEFALSYTRSMHERNLSSLSEKOPHELFAGGMASLOKSEVLGKRWL
LFLVIGLCKSKGETRRLJEQKGYYINNVPI ANEHSVCEEQDI CYGHYVLLAQGKKRKLVL
YLN

CPn_0362 407843 407055
flia/rpsD-sigma 28/WhiG Family
LDKKKFVKTQQTQNJIEVWNFYWETQEIEYRDSLIEFYLPLVKSVVHRLISGMPSHVKTE
DLYASGVEGLVRAVERYNPERSRRFEGYAVFLIKAAIIDDLRKQDWVPRSVHQKANKLSG
AMDSLRQSLGKEPYDLELCEYLNISQOELSGWFVSARPALIVSLNEEWPSQSDEGAGMAL
EERIPDERAETGYDVVDKQEFSLCLANAIQELEEKERKVMALYYYEELVLKEIGKVLGVS
ESRVSQIHSKALLKLRAALSAFR

CPn_0363 409700 407943
flha-Flage lar Secretion Protein
eAVFVSGKKDSVRGMIFVPLSILVLIFLPLPQILLDFGLCISFALSLLTVCWVFTLNSSN
SAKLFPPFFLYLCLLRIGLNLASTRWIVSSGTASSLIVSLGSFFSLGSLWAATFACLLLF
FWNFLMVSKGSERIAEVRSRFFLEALPAKOMALDSDLVSGRASYKAVKKOKNALIEEGDF
FSAMEGVFRFVKGDAIISCILLLVNIVVSVTCLYYYTSGYALEGMWFTVLGDALVSOVPALL
TSCAAATLISKIDKEESLLINTLFEYYKOLRGHFRVVSLLIFSLCCIPSSPKFPIVLLASL
LWLAYRÆEPASEDSCIERAFSYVEGACPKEQESOFYQVYRAASEEVFEDLGVRLFVLTS
LRIEERFALRVFGONVYLDEMTPEAVLPFLRNIAHEALNAEVVOKYLEESERVFGIAVED
LYPKKYŠLSSLVVLSRLLVRERVSLKLFPKILEAVAVYONSGDSLEILAEKVRKSLGYWI
GRSLWDQKQTLEVITIDFHVEELINSSYSKSNPVWQENVIRRVDSLLERSVFKDFRAIVT
SCETKFEMKWLDPHFPDLLVLSHDELPKEIPISFLGIVSDEVLVP 409700 407943 CPn_0363

410238 CPn/0364 409954 Tely4-Ferredoxin IV
KENSMAKUVITSDDEQQEFELEDNSEIAEPCESMGIPFACTEGVCGTCVIEVLEGRENLS
EPTEPEYDFLGEPEDSNERLACQCRIKGGCVKVTF

Pn_0365 410498 411544
No robust homolog present in Genebank/EMBL as of 11/7/98
FKCTQVNSLIMATISPISLTVDHPLVDTKKKSCSNFDKIQSRILLITAIFAVLVTIGTLL
IGLLLNIPVIYFLTGISFIAVVLSNFILYKRATTLLKPRACCKHKEIKPKRVSTNLQYSS
ISIAINRSKENWEHQPKDLQNLPAFSALLTDNPYEIWKAKHSLFSLVSLLPGGNPEHLLI
SASENLGKTLLIEETSQNAPISSYVDTTPSPKSLLNEAIQETRVEINTELPAGDSGERLY
WQPDFRGRVFLPQIPTTPEAIYQYYYALYVTYIQTAINTNTQIIQIPLYSLREHLYSREL
PPQSRMQQSLAMITAVKYMAELHPEYPLTIACVERSLAQLPQESIEDLS 410498

CPn_0366 411976 412440
No robust homolog present in Genebank/EMBL as of 11/7/98
MGYLPVSATDVLFSFRAPLINSANTONOKLIELKGQAESSPRTITSVILEVLLVIGC
CLIVLSLIAIRPALOFILETGHPANIAVLAVSGTILLVAVIILFCFLAAVPFAAKKTYKY
VKTVDDYASWHSHQQTPTLGTIFSGIVYAESQAQL

413078 413836 No robust homolog present in Genebank/EMBL as of 11/7/98
SFPLNRYFMTKTTSIPDVHENQSHLSVDERLISESPVLTKKEVIAKIIKLTALILALAIA
VGTAVVAGVLGMPLMAIATGAALLAAVVLSCLLLRRREPSKPTEELLGPQKHVPKDIAAQ VQPSVPLDYQKLLRNEWTLVNTLSEINISWTLQDPNQRYYVWEHQGAPITLVATTGDIAK PRLKYDSGRVMIVNAANSNMQSGGAGTNAALSAATHPTCWNNTRTSGGKINTGKGLSVGEC RSAPWINRDWINK

CPn_0368 413766 414107
No robust homolog present in Genebank/EMBL as of 11/7/98
TLAKDYLWVNAAQHPGSIETGRINDTNDGEAHFLAQLLGPKYEGELKAHPEKLSNVIKKA
YLNCFDEALNNQATVVQVPLISSSIYSPGGKLELEPVNQTKPNSSAYKLYHIRT

414345 415562 CT058 hypothetical protein_2 NIMTDSNPLPSYTDASLYRTPAKHSYPIRLPLNRTDRIEKILKIVTLTLALACALGFSIA AGILAMPIFSAVVITLAIAAVSLYSLLKKPKLYEILPOIEPESEOSSLSPSPOPPEOOD LPLOIPPLPDPESLPEVSLADLTTPPEELTAITVTPGYEALLEONWOLLPSLAAVDPSFT TETPQQPCFIWKLKDSKLIFISTSGDIAVPRIKTGGRYMIVNAANENISREGGGTNKALS LATSLOCWNASRLFRAHGRSGSQLOPGECRSAKWENSDHTSNDHVPGKAHFLAOLLGPEA AKCNNDPKQAFEVSKKAFHNLFQEAEIIGVDVIQLPLIGCNLFAPSRLLNLGKTRAEWIE AIKLALITSLQDFGWEQDNQEEQKIIILTDKDQPPIIPPRFDLTTP

415755 416912 CPH_0370 415755 416912
CTOSB hypothetical protein_3
KRIFFKLFVFYLKSFMSTTEPNLTNVNLTMLISSESMPTQLASHKLKGLDLVAFILIIGI
AVSSGTAAIILGIFLLFILTALAVLAFSILLYFLLRERKSPISVTHQPTPIIKDTDLPPV
PPLALTEVPTEAVLEEPPLFSPRTHOTILDGENWDIPIDLQANTOMPFIAADNQTGYAMHL
KNSNLTLISTLGFIEKPRYKTGGIVMIVNAATPNMANNVKTTSLALAKATSVRCWENSKK
SPDPLRSKQPLQLJECKSAKWENLHGTTNAGKAGLPQFLGQLLGPKASDYNYNPNDAFTF
CRQAYLNCLMEAKRRYTTV\QLPLLGSHFPGSPKDEETTSLRLQWIDGVKLALIDALQTF
GSEAENQNQPWVIILTTLARHPLITP

CPn_0371 417141 417503 No robust homolog present in Genebank/EMBL as of 1177798 KTMPVSSAPLPTSHRPSSONLILMERENKALKARHQDETTKETKLLVKILVATLVIEVLG LIAAFFIPSTPPICLITICALILTTY/CVLLL/TIKLALVHKTEXTTARQQIKRKLSSKSI

417651 418061 CPn_0372 CPR_0372
No robust homolog present in Genebank/EMBL as of LL77798
NYRACHRITORINSSEVYTOTOSASP/EQTE/LEFLERESSEGROIKIKIAFAASTALLLIN
TEVOGIVAIAMIEVATSVGAYFT/IGFLELLS/LLLAIMLISMYKITHPSQNTPISN

418356 420218 CPn_0371

GCPE
NSEIFEIFMTLITPAINSSRRKTHTVRIGNLYIGSDHSIKTQSMTTTLTTDIDSTVEQIY
ALAEHNCDIVRVTVQCIKEAQACEKIKERLIALGLNIPUVADIHFFPQAMLVADFADKV
RINPGNYIDKRNMFKGTKIYTEASYAQSLLRLEEKFAPLVEKCKRLGKAMRIGVNHGSLS
ERIMQKYGDTIEGMVASAIEYIAVCEKLNYRDVVFSMKSSNPKIMVTAYRQLAKDLDARG ERIMOKYGOTTEGMVASALEYLAVCEKLMYRDVYS SHASSNEKHAVYAR QLAXDLUANG WLYPLHLGYTEAGMCYDG ILKSAVG IGTLLAEGLGDTIRCSUTGCPTTEIPYCDSLLRHING CHARLESTER GERMAND IN AMERIKAKTUM ID CHYPLHULYFHHLTOFTPEELLEH GREINFERFARTTUM WYZHAFLARATTUM ID CHYPHHIROYFILYFHINEEL WODDANG VHQAPFVHFHASDPFIHTSRDFFEKQGHQKPTKLVFSRDFINKEEAALSI ATEFGALLA DELGEAVYLDLENLEYLQDVLKTAFTTLONAGVRLVKTEYISC PMCGRTLFDLEEVTTRIR KRTQHLPGLKIATMGCIVNGPGEMADADFGFVGSKTGMIDLYVKHTCVKAHIPMEDAEEE LIRLLOEHGVWKDPEETKLTV

CPn_0374 420209 420961
CT056 hypothetical protein
vDsmtLsfhTheLnywffEEFFOCLPIRHGVFSKQKDAEGTVFAAKNPEIASALQSPKYCD
LHQRHGTSVRCVTPTSPTYQPADGLCTQSPLLSLHIRHSDCQAAIFYDREHHAIANVHSG
WRGLLGNIYAVTVGTMKKLFHTKPQDLFVAIGPSIGPDYAIYPDYATLFPRSFLPFMNPK
NHFDLRAIARKQLTNLGISKDRIFISDLCTYTEHDAFFSSRYLAHHPDPNLTGQHSKNRN

421615 421112 CPT_U3/5
No robust homolog present in Genebank/EMBL as of 11/7/98
RLSMKLGASTNHKVHEPVKPKKAQLAEIEANKTQATEGTLRSKSLALQIARAVLYILFAA.
LMLAAGITFVTFLALGFPLIQAYSIAGIITLVGLAIGUVLLISLLPKEDEEADALSRNA
LHELTIIVIEQOPITFKPEIPYSYLTKLALLTSLFLTLRRSSSQRKTH

CPn_0376 421680 422294

No robust homolog present in Genebank/EMBL as of 11/7/98
FKVVTAKAPNLTEIRDHGARVPSLFLLSPETSHWKGDKEVSAPLKOLODLIGEEQWEANK
TKMNSRKKAGGWAIFNSPTGVSSTLVLAWTPWYDKDVODILERKDPMSSSLSEKDSK
EFLKNLFVDLLENGFTSVHIHAEEAFTPLDHTGKPHFKRDNVYLPGKLIGALNEAAVQAN VSADTQFTLFLTQDECNPFHDKKRG

CPn_0377 423441 422347

sucB-Dihydrolipoamide Succinyltransferase
IMTFEVRIPNIABSISEVTVASILVYEEGALIOENQCLLEIESDKVNQLIYAPVSGRIFWE
VŠEČDVVPVGGVVGKI EPAGEGEELDDSQSKETIEAEIICFPQSGVRQSPPENKTFIPLR
DQMEDGSQGLSAGDRGETRERMTSIRKTISRRLLSALHESAMLTTFNEVYMTPLFHLRKE
KQEEFLSRYGVKLGFMSFFVKAVLEALKAYPRVNAYIDGEEIVYRHYYDISIAVGIDRGL
VYEVIRCDKLSNGEIEQKLADLALRAREGLLAIAELEGGGFTITNGGVYGSLLSTPIIN
PPÖVGILGMHKIEKRPVVLDNEIVIADMMYVALSYDHRLIDGKEAVGFLVKVKEGLENPA
SÄLÄDL

CEMA-0378 426195 423445

SucA-0xoglutarate Dehydrogenase
IVFTERNYFMDSFYGOVYSSOMDMIESHYQFRMHETLDPSWKYFFEGYQLGQAASPSE
ASYKISGNETIAMLQQKSQFLCTIYRYYGYLQSISTLAPTIDSRFTQEKIAKIDLDEQ
VPSAGLLPKAQVSVRELIEALKKCYGGSLTLETLTCTPELOEFWNILMEKRQVERFABOL
LBSYKDLCKATFFEEPLOJIKFTOGKFSLEGGETLVPMLEHLWHGSALGISNYVLGMAH
RGRENVLTNVLGKPYRYVFMEFEDDPAARGLESVGDVKYHKGYVLKSHQKDRETTFVMLP
NASHLESVDPIVEGVVAALQHQGHAGKEQSSLAILVHGDAAFSCGGVYETLQLSRYFGY
STEGTLHIVNNYJGFTAVPRESRSTPYCTDIAKHGIFVFRVNSEDVVACIEAIEYALQ
VREBFSCDVIIDLCCYRKYGHWESDDPSVTAPLLXDQIKRKKSIRELFRQVLLBGQFADI
STETLASLEFFLOSSLINDEFOUNKGCHDPSPEPKEKEHUNDEINWSELLHHODYSLDRET VREFFSCDVIIDLCCYRKYGHNESDDPSVTAPLLYDQIRRKKSIRELFRQYLLBOQFADI
SEFTLASIEKEIQESLNREFQVLKGTDPEPFPKKECHHCDRLNNGELIHDCDVSLDRET
LEMSSRLCGFPDNFHPHRKIKTLLEKRNKMAEGGVGYDWAMAEELAFASLLIBGYNLRL
SGODSIRGTFSQRHLWWSDTVTDDTYSPLYHLSAEQGSVEMYNSPLSEYAILGFEYGYAQ
QAERTLVLWEAQFGDFANGAQIIFDQYISSGIQKMDLHSDIVLLLPHGYECQGPEHSSR
IEBYLGLAANWNFQVVLPSTPVQYFRILREHAKRDLSDFLVIFTFKLLLRYPQCVSSIEE
FTEPGFRAILEDADPNYDASILVLCSGKIYYDYAEMLPQDRRKDFSCLRIESLYPLALE
DLYSLIDKYSHLKHFYWLQEESKNMGAYDYMFMALQDILPEKLLYIGRPRSSSTASGSAK
IGBEFLYDTMFTIFSIR LSROELVTCMETLFSLR

CP 0379 CPT-0379 426268 426765
CT053 hypothetical protein
KNKKMLCTCSRIODCHPWMKSERLKKLESELHDLTQWMQLGLVPKKEISRHQEEIRILEH
KIYEEKERLQLLKENGEIEEYVTPRRSPAKTVYPDGPSMSDIEFVEPTETEIDIDPGETV ELELTDEGREDGAVEVDYSHEDDEDPFSDRNRWRRGGIIDPDANEW

CPn_0380 426671 427876
hemN-Coproporphyrinogen III Oxidase
KSTIPTKTMKTLSAIAIAGDAVVSLIPMLMNGKAPLALYIHIPFCTKKCRYCSFYTIPYK
SESVSLYCNAVIOGEGKELAPIQETHFIETVFFGGGTPSLVSPLDLKRILKELAPHAREI
TLEANPENLTVSYLRQLQETPINRISVGVQTFDDSILQLLGRTHSSSAAITALQECQNHG
FSNLSIDLIYGLPTQSLEIFLSDLHQALTLPITHISLYNLTIDPHTSYKHHKILVPTIA
GEEILARBSLLAENLLLSQGFQRYELASVAKPDYPAKHNLYWTDRPFGGVSASOYLH
GERSKNYSHISHYLRAVRKNLPTQETSEILPKKERIKEALALRLRLLEGAJLAEFPSTLI SMLTQDVKLQNLFSVHGQCLALNRQGRLFHDTIAEEIMGYSF

CPn 0381 429836 428037 CT126 similarity
SLPNKFRALMTAPTESRSSPPTLLEETEPLSPNPIPADIQIPRITISPPSLDVSTVASSA

430752 430036 C17E_0382

430752

43036

430752

43036

PVTLYLLPNTLSTRAVETLESVIGELVIRLDGLIVESDEGGRAFLSLWKIPEVHKEPLAI

LUBHARLUKAWDEYLEFIVKHGENGLIGDAGLCCIADGAGLVRRARADJIPVOARSGP

CSITTLAIMUSGLERGOFFELVYLDGDYRERWSIKKKAYTSKEVSTSVIETVRNYTFE

ULLDFLPGYAELGVAUDLBGPSELVLTROVOSWRTTEDLGGVKOSITKVPTIFLFHIPN

PPHITAVYOLINKSAEENWSSGDVLAVGOTLNFILCAFVLFYNLWFFVKSVLRHSRRRRR

CPL_0383

CT047 hypothetical but the control of the

13,3497 1 hctB-Histone-like Protein 2
vitclingikmigaokkosgkkyasravrkpakkvaakrtvkkatvrktavkkpavrkta
akktvakkttakrtvrktvakkpavkkvaakrvvkkttakravrktvakkpavrkt
ttvakgspkkaaacalachknykhtssckrvcsstatrkhgsksrvrtahgwrhqlikmm

434042 432522 CPn_0385 4 #4042 432522

pepA_Leucyl Aminopoptidase A

FLVIKGEFVVLFHAQASGANRVKADAIVLPFWHFKDAKNAASFEAEFEPSYLPALENFQG

KTGEIELLYSSPKAKEKKIVLIGLGKNEELTSDVVFQTYATLTRVLRKAKCSTVNIILPR

ISELRLSAEEFLVGLSSGILSLNYDYPRYNKVDDRILETPLSKVTVIGIVPKMADAIFRKE

AAIFEGVYLTRDLNNRADEITPKKLAEVALNLGKEFPSIDTKVLGKDAIAKEKMGLLLA

VSKGSCVDPHFIVVRYQGRPKSKDHTVLIGKGVTFDSGLDLKPGKSMLTMKEDMAGGAT

VLGILSALAVLELPJWTGIIPATENAIDGASYKMGDVVVGMSGLSVEIGSTDAEGRLIL

ADAITYALKYCKPTRIIDFATLTGAMVVSLGEEVAGFFSNNDVLAEDLLEASAETSEPLW

RLPLVKKYDKTLHSDIADMKNLGSNRAGAITAALFLQRFLEESSVAWAHLDIAGTAYHEK

PEDDDVBLYVASGFG/WSILVYIFNSLSK EEDRYPKYASGFOVRSILYYLENSLSK

434543 434046 CPn_0386 CPT_U386 434543 434046
ssb-SS DNA Binding Protein
KSKGYLMMFGAFACYLGADPEERMTSKCKRVITLRLGVKTRVGMKDETVWCKCNIWHNRY
DKMLPYLKKØSGVIVAGDISVESYMSKDGSPQSSLVISVDSLKFSPFGRNEGSRSPSLED
NHQQVGYESVSVGFEGEALDAEAIKDKDMYAGYGQEQQYVCEDVPF

CPn 038 435229 434699 CPT_U38/
TOTAL Mypothetical protein
NNNNLLGDSLMSRQNAEENLRNFAKELKLPDVAFDQNNTCILFVDGEFSLHLTYEEHSD
RLYVYXPLLDGLPDMYQRKLALYEKLLEGSMLGGGMGGGVGVATKEQLILMHCVLDMKY
AETNLKAFAQLFIETVVKWRTVCADICAGREESVDTMPQMPQGGGMQPPPTGIRA

CPN_0388 435323 437320
glgk-Glycogen Hydrolase (debranching)
stylekvssypsyplplgaskispnryrfalyasgatevilaltdensevievplypdthr
tdalwhieiegisogssyafrvhgpkkhgmgysfkeyladpyaknihsposfgsrkkogd
yafcylkeepfpwodoplhlpkeemityehhvrsftossssrvhapgtflgiiekidhl
kklginavellpifefdetahpfrnskfpylcnymgyaplnffspcrryayasdpcapsr
Efktlvktlhqegievildvvfnhtglogttcslpwidtpsyyildaoghfthysogont
Lntnrapttomildilrwweemhvdgfrfblasvfsrgpsgsplofapvleaisfdpll
astkiiappwoagglyogypflsprwsemppyrdnykaflngonligtfasrisgs
odiyphgsptnsinyvschdoftlcdtytynkhkheangebnbgtdanysynfgtbske
Edpgilevrerolnnfflthvsogipmiosgdeyahtaegnnnrwaldsnanyflwdol
takptlwhflcdliafrkkyktlfnrgflsnkeiswvdamghpmtwrpcnflafkikspk
hvyvsphydgodolatlpkasshelpyojvaesoogfyponvaffvslophttlinis AHVYVAFHVGAQDQLATLPKASSNFLPYQIVAESQQGFVPQNVATPTVSLQPHTTLIAIS NAKEVT

Pn_0389 438254 437319

CT041 hypothetical protein

TVFNFKRFYQKDSQRQNGNTTCLRFFKKTCKELIEFRRTVKLLKNVLLGLFFSMSISGF
SEVKVSDTFVKQDTVVEPKIRVLLSNESTTALIEAKGPYRIYGDNVLLDTAIQCQRCVVH
ALYBGIRWGEFYPGGCCLKIEFVDDTASLFFNGIQYQGSLYVHRKDNHGIMVSNEVTIED
YLKSVLSIKYLEELDKEALSACIILERTALYEKLLARNPQNFWHVKAEEEGYAGFGVTKQ
FYGVEEAIDWTARLVVDSPQGLIIDAQGLLQSNVDRLAIEGFNARQILEKFYKDVDFVVI

CPn_0390 439171 438134
ruvB-Holliday Junction Helicase
RKSDREGSYMTHOVAVLHODKKFDVSLRPKGLEEFYGGHHLKERLDLFLCAALQRGEVPG
HCLFFGPPGLKTSLAHIVAYTVGGLVLASGPOLIKPSDLLGLLTSLQEGDVFFIDEIH
RMGKVAEEYLYSAMEDFKVDITIDSGPGARSVRVDLAPFTLVGATTRSGMLSEPLRARFA FSARLSYYSDODLKEILVRSSHLLGIEADSSALLEIAKRSRGTPRLANHLLRWVRDFAQI REGNCINGDVAEKALAMLLIDDWGLNEIDIKLLTTIIDYYQGGPVGIKTLSVAVGEDIKT LEDVYEPFLILKGFIKKTPRGRMVTQLAYDHLKRHAKNLLSLGEGQ

439701 439510 No robust homolog present in Genebank/EMBL as of 11/7/98 KDQLYKQEKPIPKATILSRNLEVMLDNPKGKRQTLFLGRTSGRSALYSYSRRILVLLNAF

439814 440383 dcd-dcTP Deaminase
MSIKEDKWIREMALNADMIHPFVNGOVNVNEETGEKLISYGLSSYGYDLRLSREFKVFTN
VYNSVVDPKCYFEDIFISITDDVCIVPPNSFALARSVEYFRIPRNVLTMCIGKSTYARCG
IIVNVTPFEPEWEGHVTIEISNTTPLPAKIYANEGIAQVLFFESSTTCEVSYADRKGKYQ KQQGITVPCV

CPn_0393 440229 440723 CT039 hypothetical protein Ketlehecorkftlakgiprgysisivpparfimotekesiksnkaspylvskvsvrkkn WGFRLLEEVMIKSWWVIFSILIGGFV/DRAIOELRTEELRLQSKVSSLCODILSAQEKQR QLQLHLQHWQDSAATEAALTORLGLTPKGYKKLCVSPKQQSENKD

440727 441968 CPD_0194 440727 441938
ttyC-CBS Domain protein (Hemolysin Homolog)
KETMIPTMLMFFI ICFTLCOOFIGLOC/TALECLERBLICHYKROKSKKOORVATLLLHPH
HELITELFCDIGLATATONCEATLEGDAASMMETVALHELATTLITLEELLEKAVALPENTO
TASSVAPLILAVTKIFKELLHMGIVGIGTVVVVMIGSKGOJOT IQUOELKEVLOSKOPOV
NOBESPELLAVYLSEKELLHMGIVGIGTVVVVMIGSKGOJOT IQUOELKEVLOSKOPOV
NOBESPELLAVYLSEKELLHMGIVGIGTVVVMIGSKGOJOT IQUOELKEVLOSKOPOT ICHDN
LONLIGITARSILLHIDKPLOGSDDLIGHEKEPTVMHETTISAKMALCOMAAEDETISMIT
DEYGSJEGGLIVOBEDFETVAGETVDGPENKLISMINILIFOVIDAAFMIRTRKVYTKKLYD
PYNNNIATICSWLIEQIGT IPPTYMKLISMINILIFOVIDAAFMIRTRKVYTKKLYD CPn_0195

441955

443175

CT257 hypothetical protein
CNCMTNSALFWIGOVILLCIVLOOFYSMMEMACVSFNRV

PYRLFGTVMIGOVILLQVGSESSRNCYRALGITPDVAPFTQIFIVVIFAELLPLTISRKI
PEKLALMGAPILYYSHYIFYPLIQLIGSLTEGLYYLLNIRKEKLNSTLSRDEFQKALETH
HEEQDFNTIATNIFSLSATCADQVCQPLEQVTMLPSSANVKOFCRTIKNTDINFIPVYHK
ARKNVIGIAHPKDFVNKALDEPLINNLHSPWFITAKSKLIRILKEFRDNRSSVAVVLNAS
GEPIGILSLNAIFKILFNTTNIAHLKPKTISVIERTFFCNSSIKDLQKELDIQFPQYPVE
TLAGLUTAGLEDFAFFYTGVIIHRILLEFFKEMALGGIKTYGIKNILL

CPn_0396 444159 443241
yhf0-NifS-related protein
YSMIYLDNNAMTPPERGLLEFLOKTFLIEGTYANPSSVHQLGKKSRQLVLEASHWMQKVL
SFOGRVLYTSGATESLNLAIASLPKDSHVITSGSEHPAILEPLKHSSLSVSYLNPEEGRC
VLTIEQIERAVTPKTSAIILGWNSETGAKADIAAIAHFAQERQLQFIVDATANVGKERI
VLPSGVTMAAFSGHKFALASGIGALLVSFGVKLHPOLMGGGOGGGLRAGTENLGIASLL
YIFKYLDLHQERISQEILTHRNGFEKAIKARIPDVHIHCADQPRANNVSAIAFPPLEGEV
LQIALDIEGVACGYGSACSSGATAPFKSLVSMGVDEELTLATLRFSFSHLLLQEDVERAV
GIIEKVVERLKNS

CPn_0397 445124 444381
PP2C phosphatase family
EHFVDFDYFGLSDIGRVRANEDFWQVNLMSQVVAIADGVGGRLGGDIASQEAVTSLMEL
IDEQQSKLMGYGDDQYKETLKKILLEVNGVVYEHGQMEEHLQGMGTTLSFIQFRKDRAWL
FHVGDSRIYRIREGELRRLTEDHSLENQLKNRYGLPKQSDKVYSYRHILTNVLGSRPYVM
PDIRNLPCEKEDLYCLCSDGLTNMVPDIDIRDILNQPATLEERGNALISLANTRGGDDNA
TVVLVRIQ

CPn_0398 445518 445700
No robust homolog present in Genebank/EMBL as of 11/7/98
IEELPMQIENSSILFAEVVMKWFIFSVISAPVVFLPGCTLIPKEKVTKVPSQLWSESLSQ

CPI_0399 445759 446523
CT253 hypothetical protein
YKLMRVLNGKSLNCESIDLKSKNFPRARIFCKISNLRTVTMRKMLVLLASLGLLSPTLSS
CTHLGSSGSYHPKLYTSGSKTKGVIAMLPVFHRPGKSLEPLPMNLQGEFTEEISKRFYAS
EKVFLIKNNASPOTVSOFYAPIANRLPETIIEOFLPAEFIVATELLEQKTGKEAGVDSVT
ASVRVRVFDIRHKIALIYQEIIECSQPLTTLVNDYHRYGWNSKHFDSTPMGLMHSRLFR
EVVARVEGYVCANYS

CĒRĪ0400 446527 447306
CTĪ254 hypothetical protein
SKĒMSKFILLISLGVAALASKNIFFINPAPSGKTPLKLRQVLFGGALLVFSSLVALSVSSQ
TAĒLLSTMTGISLAFĀFLFYLLFIPKDITRAILFSGERPVKTSWRALGSAIRMWIIIIPV
TĢĒJGIMMSKFLTLVLPTQĒIHTQEVTQEVONSLPITCHYISMILNLGVLTPFGEEVFFR
GILGTFLKNKMTRIAAVLCSSIIFSFIHIEHSLGSWVFVPVLFVFSLSAGFLYEKDRHIL
SPĪĀLHGLFNLTSLLFLGIK

CENTO401 447884 447495
CEATS hypothetical protein
MEDIAFSKLIGTVRAMVVEGRCPWSLQQSLVSMVEHILGECQEFHEAVLQGKTVQEVGSE
AGENTLTLVLILCFLLEREGVLASEDVANEAMEKLRRRAPYIFAEDYKPVSIEEADRLWEL
AGENTLTVLICFLEREGVLASEDVANEAMEKLRRRAPYIFAEDYKPVSIEEADRLWEL

Chrio403 449009 449710
ycer_predicted pseudouridine synthetase family
NFMGLSNDKRAALOYFMENFSWLATOVSRLSSFLRSQLPNHSKQEILASIRQHRORVNGF
IERFESYKVOYGDRVSLSIIPSTKQOPSILWEDDYSIIYEKPPHLTTEQMAHMTRFFTVH
RLDKGTSGCLLMGKSKQAATELMKLFKQRKIHKQYIAFVFGHFKKKFGTVKSYYAPVYRR
CGAVIFGAAGPSQGEPIKSAYKWDCWVILLSEMSTTDLKNSLPRSSALSSMLTP

CPn_0404 450962 449871
No robust homolog present in Genebank/EMBL as of 11/7/98
ELEALEQKYGKAVLLIALSELGIDTMSLLSGHRLEGFPPIAEVMAACDROSMDFCEILKS
QSMDLWADAASCVUGLLODPFWSTAIASGIAKSSLQETEFECESKYMVLSSWGEQGAQVC
SPFNLERICMSFPSLKVFSLKKNGCENMGIQLSASCMNLLMSIFFVATWGGSTPIWITKE
NLMALVALVLSHYQCYFVPATGDPORGNILGNPEVNAILARGMGNRVVLERKRGGESSSS
RYLELAARCFENSLTKTSLLSDANNVQERDKCLLOMSTSLMHTAGLNLORPPVPTPSGVT
AHPOPQPDPVVTSQPSLLGARERSPVSSRGRFPVVLPLSVISPRSHFGRVERRDLEDEEE
EVMF

CPn_0405 451814 450966
CT105 hypothetical protein
NIQTSHSRVLLKKFSKEFTIRTYRSLGFTDYLCGCLTNPLCKFPSPQNPQVVTIAPSSTT
PQAVSSAVQGFLQTGGAASSTATTTTASGASALGLSPDQVQALLTNLLNVCQPSVGQPST
SACTSGASSSSASMQQQLLQLILDKTTGSGGSSVSSEQLQQLLSLVSOMTTSQGSGGTQ
ACQAASVLLNLLSATGSAAANPLGTAASLAQIIYAAVTSPQAKKTSEFCYNYCGETCQGN
CGCPTCGCPDQQCGCGGFGRFFCGWKNCCGIGEGSQEPAIPL

CPn_0406 451960 452865
table-Enoyl-Acyl-Carrier Protection Reductions
GGMLK (DLM)KVAPVAGIGDDGGYGWGIAKLLAEAGATTI (VCTWVP LYK IFGQSWELGK
KNESRKIJANOTELLETAK LYPMDAGFDGPEDVPED IA ØNKRYKG LYGPP I SEVAEQVKKDF
GILD I LVIIGLANDPE GKIJLLETGRKOYLAALBAJY/GFVSILDHEGG IMBROGST IGLT
YLAMBRAVPCYGKIMSIAKALESDFOKTLAMEAGIA WG IR KWYT ENAPLASRAGKA IGFT
EKMVDYYGEMGAFTER

CFm_0407 453757 45285 HAD supertamily bydrolase/phospyatase NYGDAMEKLLVTDID/TITT DKKY/YERLYALHQA/WKLFFLTOR/Y/KYAAPLFSD FDAPYLLCCONGASVWG/TT. YSKSY/PSDLLCILQDCMEGATALFSVESGAPYGDHY YRESPTPIAQDLHEYVDRY/PNAKERFILFETRSKLKDYX/FPSAAKVFGLABEVIRI OKELEROBALTSVATHTLMRWFFDFPY/AILFLTDKSVSKCKALDRVVNILYDGKKPFWIA SGDANDLDLIERGDFKIVMSSAPEBYHVHADFLAPPADKNGILSAWEAGVRYYDDLMSL

CPn_0408 454090 454581
CP102 hypothetical protein
DPVLIMPERIOR OF EACH PTTORE CONTINUED SMKLEMERMICOLILLICITES
BYWARENEYS BETWEEN CONTINUE AS FOLLOW AS TOPEN OF THE BYWAREN AS SETTING FOR THE BYWAREN AS SETTING

CPn_0409 454645 455127
CT260 hypothetical pfotein
MTTWTLNONNLTKFLKSSDEPFLERESGLTYINIQANGNELPLFFVIRSEGEILQLICY
LPYQLHESHKASTARLLHILARDIDIPGFGMDEEQGLIFYRLVLPCLNGEIHDTLLRIYI
DTIKLVCDSFSHAIGLISSGNMNLDELRRQALQEQQEKRNE

CPn_0410 455087 455833
dnao-DNA Pol III Epsilon Chain
DVRLFKSNKKNYMSSØTHDVLIFYDTETTGTCIERDRIIEIAAYNSVTDESFLTYVNPEI
PIPDEASKIHGITTDAVLSAPKFPEAYEGFRKFCGEDSILVAHNNDGFDFPLLGKECRRH
SLEPLTNRTIDSIMAGKYRPDLPKHNLQYLRQVYGFAENQAHRALDDVVILHKVFTSLI
GDLPPQQVLDLLODSYHPKVFKMPFGKYKGQPLVDIPKSYFEWLENQGALDKPENKDIKA
AIALLHQPT

CPn_0411 455794 456609
CT262 hypothetical protein
RHOSRYSSITSTDMILTAAFSPCPNDIFIFRSFLKDPOFRPLLNOVTIADIETLNTLALO
RRLSLMKNSAALFFLVSDYYNLMDVGNTLGYNSGPIVLSLDPECSLDTLATPGEMTTAHA
LCKLYYPKKLIPMPYDKILSAILGSKVDGGALIHEERFSYDLOLTLRADFGELWRRKT
FPLPLGCIAIAKYVPMATVDALTAALRKSLICSLKDPITAGAKAVEYSKNKNVTVIHRFI
GTYINKEPFQLSKTGKKALHHLWKANECCOYT

CPn_0/12 456515 457246
CT263 hypothetical procein
EPISYKKPPNYLKLCKKLYICSGRPMNAVNTPKKILCIVADYREISPLIEOLDFTOINEH
LYSYRCTDYHLDLYIVNVMGSTAVLNALOSYCOAYTDYDLWINPGFVGACSPEIPLGCY
IIEKIANLITTDTPPVLSEDPPYIFDALPDSLPKSSLVTSPVLYHYGFHKTFKLLDMEGYA
IASOAAEHHIPCSFLKITSDYTVPGDCPFSRLEEVSOKLTOTLVELLPELMERAIPPKLL
LPCP

Pn_0413 459209 457227

msba-transport ATP Binding Protein

VFMKLLLKAVLRKNHLVILGCSLLAILGLTFSSQMEIFSLGMIAKTGPDAFLLFGRKES
GKLVKVSELSQKDILENMOAISKDSETLTVSDATTVIAEHGKSTASLTSKLSKVRNYIDD

VSRFRGLAIFLICVATFRAVTLFFORFLGGVVAIRVSRDLRGOPYFKALQOLPMTFFHDD

IGNLSNRVMTDSASIALAVNSLMINYIQAPITFILTLGVCLSISWKFSILICVAFPIFIL

PIVVIARKIKNLAKRIQKSQDSFSSVLVYDFLAGVMTVKVFRTEKFAFTKYCEHNNKISAL

EEKSAAYGLLPRFLLHTIASLFFAFVVVIGIYKFAIPPEELIVFCGLLYLIYDPIKKFG

EEKSAAYGLLPRFLHTIASLFFAFVVVIGIYKFAIPPEELIVFCGLLYLIYDPIKKFG

EEKSAAYGLLPRFLHTIASLFFAFVVVIGIYKFAIPPEELIVFCGLLYLIYDPIKKFG

EEKSAAYGLLPRFLHTAISLFFAFVVVIGIYKFAIPPEELIVFCGLKJYLIYDPIKKFG

EEKSAAYGLLPRFLHTAISLFFAFVVVIGIYKFAIPPEELIVFCGLKJYLIYDPIKKFG

EEKSAAYGLLPRFLHTAISLFFAFVVVIGIYKFAIPPERVFTGVGCDCHHLKNL

SFTLHKGEALGIVGPTGSGKTTLVKLLPRLYEVSGGKLIDGSLFITEYNKGSLRNHIACV

LONPFLFYDTVMNNITCGKDMEEEAVLEALKRAYADEFILKLPKGVHSVLEESGKNLSGG

QQQRLAIARALLKNASILILDEATSALDAISENYIKNIIGELKGQCTQIIIAHKLTTLEH

VDRVLYIENGQKIAEGTKEELLQTCPEFLKMWELSGTKEYNRVFVPDHKLVANPTDMAIT

T

CPn_0414 460203 459172
acca-Accoa Carboxylase/Transferase Alpha
LCLRIVCIKMILFIRGEHILMELLPHEKQVVEYEKAIAEFKEKNKKNSLLSSSEIQKLEK
RLDKLKEKIYSDLTPWERVQICRHPSRPRTVNYIEGMCEEFVELCGDRTFRDDPAVVGGF
VKIQQRFVLIGQEKGCDTASRLHRNFGMLCPEGFRKALRLGKLAEKFGLPVVFLVDTPG
AYPGLTAEERCOGWAIAKNLFELSRLATPVIIVVIGEGCSGGALGMAVGDSVAMLEHSYY
SVISPEGCASILWKDPKKNSEAASMLKWHGENLKOFGIIDTVIKEPIGGAHHDPALVYSN
VREFIIQEWLRLKDLAIEELLEKRYEKFRSIGLYETTSESGPEA

CPn_0415 461522 460221
CT266 hypothetical protein
SQTGFLPGLTLIFVIIIWCNAFLIKLCVIMGLQSRLOHCIEVSQNSNFDSQVKQFIYAC
QDKTLRGSVLKIFRYHPLLKIHDIARAVYLLMALEEGEDLGLSFLNVQQYPSGAVELFSC
GGFWKGLPYPAEHAEFGLLLQIAEFYEESQAYVSKMSHFQQALFDHQGSVFPSLMSQE
NSRLLKEKTTLSQSFLFQLGMQIHFEYSLEDPALGFWMQPTRSSSAFVAASGCQSSIGAY
SSGDVGVIAYQPCSGDISDCYYFGCCGIAKEFVCQKSHQTTEISFLTSTKDKPHPRNTGFS
YLROSYVHLPIRCKITISDKQYRVHAALAEATSAMTFSIFCKGKNCQVVDGPRLRSCSLD
SYKGPGNDIMILGENDAINIVSASPYMEIFALQGKEKFWNADFLINIPYKEEGVMLIFEK
KVYTSEKGRFFTKMN

CPn_0416 461871 461557 himD/ihfA-Integration Host Factor Alpha EALSNMATMTKKKLISTISQDHKIHPHHVRTVIQNFLDKMTDALVKGDRLEFRDFGVLQV VERKPKVGRNPRNAAVPIHIPARRAVKFTFCKRMKRLIETPNKHS

CPn_0417 463047 462244
amia-N-Acetylmuramoy1 Alanine Amidase
REKGMKLTKYLMTKOLRSMISRLFVRYSLEMSKQLSFFALCVLGSHPIFAQTPNPPQRVR
RSEVIFIDPGHGGKDQCTASKELHYEEKSLTLSLALTVGSYLKMETYKPQLTRSSDVYVD
LGKRVALSNRGQGDVFISIHCNHSSNAAAFGTEVYFYNGVRGSPTRNRMSEVLGKNILAA
MEKNGILKSRGLKTANFVVIRDTSMPAVLVETGFLSNSPERAALQDARYRMHVAKGIAEG
VHNFLGGPSFQKPKQNIAKIRKPQIQAN

466897 454875 CPn 0419

CPn_0419 456877 454876
pbp]: transplycolase/transpeptidase
OLFPNTNIWNIPOKKVSVFYPMSYRKRSTLIVLGVFALYALLVLRYYKIQICECDHWAAE
ALGOHEFCVRDPFRRGTFFANTTVRKGDKDLQOPFAVDITKFHLCAPLAIPECHRDEII
OGILOFIEGOTYDDLSLKLDKKSRYCKLYPLDVSYHDRLSLWMKGYATKHRLPTNALFF
ITDYGRSYPFGKLLGQVLHTLREIKDEKTGKAFPTGGMEAYFNHILEGDVGERKLLRSPL
NRLDTNRVIKLPKDTSDTYLTINPVIOTIAEEELERGVLEAKAQGGRIILMNSOTGEILA LAGYTERDITTERRETEURER FRITTEZERZIÖZERDE MKELTZA LAGARERADENG KKI FERRER HATTETERRITERRETER GEREN GEREN GAMYMA TOKOGRATZAGLARR HOSIG VAWYQOKLLALGEGRKTG I ELPSEASGLVPSPHRFH I NGSLEWSLSTPYSLAMGYN I LAT GIOMVOAYAILANGGYAVRPTLVKKIVSASGEEYHLPTKEKTRLFSEEITREVVRAMRFT TLPGGSGFRASPKHHSSAGKTGTTEKMIHGKYDKRRHIASFIGFTPVESSEGNFPPLVML VSIDDPEYGLRADGTKNYMGGRCAAPIFSRVADRTLLYLGILPDKKLRNCDEEAAALKRL YEEWNRSPKQGGTR

CPn_0420 467120 466824
CT271 hypothetical protein
KSPPMMXSRFLRLCCCLFPGSLFYFYINKQNSLTKLRLEIPCLSVRLRQLEQQNISLRF
LIDKIERPDHLMEIAALPEYQYLEYPSEESISLLSYELP

CPn_0421 468007 467108
yabC-PBP2B Family methyltransferase
EILMSERAHIPVLVEECLALFAQRPPQTFRDVTLGAGGHAYAFLEAYPSLTCYDGSDRDL
QALAIAEKRLETFODRVSFSHASFEDLANQPTPRLYDGVLADLGVSSMQLDTLSRGFSFQ
GEKEELDMRMDQTOELSASDVLNSLKEEELGRIFREYGEEPQWKSAAKAVVHFRKHKKIL
SIQDVKEALLGVFPHYRFHRKIHPLTLIFQALRVYVNGEDRQLKSLLTSAISWLAPQGRL
VIISFCSSEDRPVKWFFKEAEASGLGKVITKKVIQPTYQEVRRNPRSRSAKLRCFEKASQ

468233 468784 CT273 hypothetical protein GLAWVEIFNYSTSIYEQHASINRIVSDFRKEIQMEGISIRDVAKHAQILDMNPKPSALTS LLQTNOKSHWACFSPPNNFYKORFSPYLAPSLGSPDQQDEDIEKISSFLKVLTRGKFSY RSQITPFLSYKDKEEEEDEDPEEDDDDPRVQQKVLLKALDLGVKSTNVMIDYVISRIFQ

468788 469216 CT274 hypothetical protein
CMLDNEWKAILGWGDDELEELRISGYSFLRQGHYSKAILFFEALVILDPLSIYDHQTLGG
LYLQIGENSQALAVLDQALRWQGDHLPTLLNKTKALFCLGRIEEATAIATYLSSCPIPAI ANDAEALLMSYSKATKKNAALVR

CPTE-0424 469528 470961
dnäf-Replication Initiation Factor
SRCMEIFSPSIMGWVDCIMESFINKESGMITCHECTTWEOFLNYVKTRCSKTAFENMISP
IOVIEETOEKIRLEVPNIFVONYLLDNYKFDLCSFVPLDVHGEPALEFVVAEHKKPSAPV
ASQUESNEGISEVFEETKOFELKLALSYRFDMFIEDPSNOFVKSAAVGIAGKPGRSYNPL AGMESTED IS VELTANDELIKLAISTIKE DAY I BEFSNOF VISAAVOLAUKPERS YNFE FIRGSVOLGKTHLLHAVGHYVEEHKKINETHEI TEAFINDLYVHLKSKSYDKIMINFYKS LDLLLVDDIOFLONRONFEEEFCNTFETLINLSKOIVITSDKPPSOLKLSERIIARMEMG LVÄHVOI PDLETRVAILDHKADCKOLLI FINEMAFYLADHIYONNYOLGGAINKLTAYCRL FOKSLTETTVRETLKELFRSPTKOKISVETILKSVATVFQVKLNDLKGNSRSKDLVLARO IAMFLAKTLITDSLVAIGAAFGKTHSTVLYACKTIEHKLQNDETLKRQVNLCKNHIVG

CT276 hypothetical proteins
FRGCPMFRATGKGFFEDVOTLYEETTSSPSYSPYSRSERPETPPSLFDNPKASEARPLN
FNLTEESSLPQWSSTPRTESLLPLEEPETTLGEGVTFKGELAFERLLRIDGTFEGILVSK
GKIEIGFKGVVKADIQLQEATIEGVVEGNITVSGKVELRGGAIIKGDIQANTLCVDEGVR
ILGYLAIAGITDHSERERDL

CPn_0426 CT277 similarity 472111 471536 WYSFSLIFPKLCYGCQAPGAYFCSNCLEKLLVEDREGRCLHCFRYLGSSETRLCSQCSPS SQEGAFSLYLPSGTALSVYARACEGKRPALQFFSKSIAFELASLDETPSCIAYITSTISR KIVYEVAKLEKLLRIPLWPWLPKKRQIEKLPKGEGICFLSAYPLSQKWMQTIVGGSASPL VSTSEFLSONDQ

CPREGAZ7 472153 473715

nqr2-NADH (Ubiquinone) Dehydrogenase
AVCYYFERVEASTFLSITMLKKFINSLWKLCOODKYORFTPIVDAIDTFCYEFIETPSKP
PFIRDSVDVKRWMMLVVIALFPATFVAIWNSGLOSIVYSSORVLMEOFLHLGGFGSYLS
FVYKEIHLVPILWEGLKIFIPLLTISYVVGGTCEVLFAVVRGHKIAEGLLVTGILYPLTL
PPTIPYWMAALGIAFGIVVSKELFGGTGMNILNPALSGRAFLFFTFPAKMGDVWVGSNP
GVIKDSLMKMNSSTGKVLIDGFSQSTCLQTLNSTPPSVKRLHVDAIAANMLHIPHVPTQD
VIHSOFSLWTETHPGWVLDNLTLTQLQTFVTAPVAEGGLGLLPTOFDSAFAITDVYSGV
KFSAGNLFWGNIIGSLGETSTFACLLGAIFLIVTGIASWRTMAAFGIGAFLTGWLFKFIS
VLIVGONGAWAPARFFIPAYRQLFLGGLAFGLVFMATDDVSSPTMKLGKWIYGFFIGFMT
IVIRLINPAYPEGWMLAILLGNVFAPLLDYFAVRKYRKRGV CPn 427 IVIRLINPAYPEGVMLAILLGNVFAPLIDYFAVRKYRKRGV

474681 473719 CPI_U428 4/3/19 4/4681
inqr3-NADH (Ubiquinone) Oxidoreductase, Gamma
NMSKGSSKHTVRINOTWYIVSFILGLSLFAGVLLSTIYYVLSPIQEQAATFDRNKOMLLA
AH LLDFKGRFQIQEKKEWVPATFDKKTQLLEVATKKVSEVSYPPLELYAERFVRPLLTDA
OGKVFSFEEKNLNPIEFFEKYGESPPCOSPLPFYVILENTSPTENMSGADVAKDLSTVQ
ALIFPISGFGLWGPIHGYLGVKNDGDTVLGTAWYCGGETPGLGANITNPEWGEOFYCKKI
FLODSSCTTNFATTDLGLEVVKGSVRTTLGDSPKALSAIDG/SGATLTCNGVTEAYVQSL
AVXDOLI LNECNI MUBEUMTOE ACYROLLINFSNLTHEKKTGE

CPn 0429 474666 CPT_0429 474666 475319
mit4-NADH (Ubiquinone) Reductase 4
KENNEMTOKKOYKSYFFDPLWSNNQILIAILGICSALAVTTTVQTAITMGIAVSIVTGCS
GFFVOLLRKFTPDSVRMITQLI LIGHFVIV TOQFLKAFFFD DIKTLSVFVGLI ITNGIVM
GROEDLARHVTP IPAFLOPFASGGGVGWVLLV IGV IFELFGFGTTLMGFRI I FOFVYAGET
HPD/YQHLDIMVLAPSAFFLLGIM IWLAN IRDSKKFKR

475323 CDTPPHILE WCTOPPTING LAMAPMOUTG LOCKFOAK I GRAPLEPEVVEPTTNPLKESS 476093

SKHOPSISKARTORRSL

4764130 176151 n Genebank/EMBL as of 11/7/98 No robust homolog present KIMTTLPKYVPRSRONPOTLTFLKRYSSVLLHSENGLSYRIFAKVLAILLTSLAVAFAVT LFSCEGSQLRLCALYIGIALAICVLLTIVVYCIASKIATACKKPPSISRIEIV

r - 476514 n Friederich Pembler in Fried on twicklichterwikkerpielenwichtelstag 2 · · · - 99 density from the right E LIBERT UMDPACIEVARIO LIMLVCACWRYKLINALERTKVAHES GLFLTGSAPLULJIWIAAGGT

GSH-Glycine Cleavage System H Protein RTFRILYGTLYRTGSRKVMYSDYHVWILPVHERVVRLGLTEKMQKNLGAILHVDLPSVG SLCKEGEVLVILESSKSALEVLSPVSGEVIDINLDLVDNPQKINEAPEGEGWLAVVRLDQ DWDPSNLSLMDEE

CPn_0434 479471 477276

CT283 hypothetical protein

RPWVRIYQODLFCRUCRDPAWFFSLLSFTLRFYCLGRGWTLLSFFYKHQKKFIGIVIAVV

CVSIGVGWGRFSRKGSAESTSRRTVFTTASGKRYVEKDFMAMKKFFAHEAYPFTGNPRA
WNFINEGLLTDYFITRYGEKLFLKVYHRGEKIFSKEKAYQPYRRFDAPFISSEEVWKSS

APOLLEILKVFOGIENPISKEGFLARAKLFLEERRFPHYVLRQMLEYRROMFALPPDEAL

SRGKDLRLFGYGTIQDWFGDAYLSAAVELLIRFIDEQKKVLPRPSKQEARDDYDKAKHA
YTKISKNKEFSLGFEFVNSYFQFLEISESEFFNMYRDILLCKRALLLLQGGVSFDFOPL

TTFFVQGKDS/OVEFFRLPKEYSFKTKQELKAFEVYLKLVSLPKSDSLDVPNEILPIATI

KAKEPRLVGRAFSIDYKRVALQDLAATVPMVEVLHWQONSEHFQEILQGFPDVETCQSYK

DFOHLKPALFOKISLFTRKEILRARPERILOSLQOVPKQSOEVLLSAGKNSALPGISDGQ

QLAKVLLDWEVLDLYSQDAETYYTIIVNSSFEKEEVLPYREVLKRDLASQLLTSHGHLVD

MERLESAIRTRYPGEEGASLWQRRLWKVVENHRLGRHLEGSFSWSLDRSLKTFSRGDKEL

PQEFDRLFSMKVGDYSSVFMSPNEGPCYYQCLSHLLYDRPASVDKLFLAKSQLDEELLGS

YMERFIPQGWR 479471 477276

CPn_0435 480908 479475
Phospholipase D superfamily (uncleavable leader peptide)
GVMMSRLRFRLAALGIFFILLVPNSVSAKTIVASDKEKVGVLVYDNSVEAFQQILDCIDH AND VELCPCMTGGRTLKEMVDHLEARNDLVPELGSYIIIOPTPTDAEDOKLLKALKERH PNAFFYVFTGCPPSTSILAPNVIEMHIKLSIIDGKYCILGGTNFEEFMCTPGDEVPEKVD MPRIFYSGVRRPLAFROQDIMLRSTAFGLQLREEYHKOFAMWDYYAHHMWFIDHEOGFAG
ACPUTLEQAEETVFPGFDKHEDLVLVDSSKIRIVLGGPHDKQPNPVTQEYLKLIQGARS
SVKLAHMYFIFKDELLNALVDVSHNHGVHLSLITNGCHELSPAITGPYAMGNRINYFALL
YGKRYPLWKKWFCEKLKEYPERVSIYGFAIWETQLHKKCHIIDDEIFYIGSYNFGKKSDAF
DYESIVVIESPEVAAKANKVFNKDIGLSIPVSHGDIFSWYFHSVHHTLGHLQLTYMPA

CPn_0436 481633 480902
lplA-Lipoate Protein Ligase-Like Protein
FYVCYMKVRIVDSGKSSAASHMAKDRDLLESLQDGELILHLYEWENPCSLTYGHFMRPEK
FLLSNYADLGLDAAVRPTGGGFVFHKGDYAFSVLMSATHPSYSSSVLENYHTVNSFVAKV LEKVFRIGGMLAPEDENSSSRDSGNFCMAKTSKYDVLFGDKKIGGAAORKVQQGFLHQGS LFLSGSSSEFYQRFLKPEVLEEIIEQIQIHAFFPLGLEAADEVLQEARQQVKEAFIKLFC

CPT_0437 481810 484350

ClpC-ClpC Protease

YMPEKFYNRAKOYIKLAKKEAQRLNHNYLGTEHILLGLIKLGQGVAVNVLRNLGIDFDT
ARQEVERLIGYGPEIQVYGDPALTGRVKKSFESANEEASLLEHNYVOTEHLLLGILHQSD
SVALQVLENLHIDPREVRKEILRELETFNLQLPPSSSSSSSSRSNPSSSKSPLGHSLGS
DKNEKLSALKAYGYDLTEMVRESKLDPVIGRSSEVERLILILCRRKNNPVLIGEAGVGK
TAIVEGLAQKIILMEVPDALRKKRLITLDLALMIAGTKYRGOFEERIKAVMDEVRKHGNI
LLFIDELHTIVGAGAAEGAIDASNILKPALARGEIQCIGATTIDEVRKHEKDAALERRI
QKIVVHPPSVDETIEILRGLKKKYEEHHNVFITEEALKAAATLSDQVYHGRFLPDKAIDL
LDEAGARVRVNTMGQPTDLMKLEAEIENTKLAKEQAIGTQEVEKAAGLRDEEKKLRERLQ
SMKQEWENHKEEHQVPVDEEAVAQVVSLQTGIPSARLTEAESSKLLKLEDTLTRKKVIGON
DAVTSICRAIRRSTGIKDPNRPTGSFLFLGPTGVGKSLLAQQIAIEMFGGEDALIQVDM
SEYMEKFAATKMMGSPRGYVGHEEGGHLTEQVRRRPYCVVLFDEIEKAHPDIMDLMIQIL
EQGRLTDSFGRKVDFRHAIIITMTSNLGADLIRKSGEIGFGLKSHMDYKVIQEKIEHAMKK
HLKPEFINRLDESVIFRPLEKESLSEITHEINKLGSLKNYQMALNIPDSVISFLVTKG 481810 484350 HSPEMGARPLRRVIEQYLEDPLAELLLKESCRQEARKLRATLVENRVAFEREEEEQEAAL

CPn_0438 485455 484334
ycbF-PF-loop superfamily ATFase
NLTLEMPROVREIMQOTVIVAMSGGVDSSVVAYLFKKFTNYKVIGLFMKNWEEDSEGGLC
SSTKDYEDVERVCLQLDIPYTVJSFAKEYRERVFAFFLKEYSLGYTPNPDILCNREIKFD
LLOKKVZELGGDYLATGHYCRLNTELQETQLIRGCDPOKDQSYFLSGTPKSALHNVLFPL
GEMNKTE/RAIAAQAALPTAEKKDSTGICFIGKRPFKEFLEKFLPNKTGNVIDWDTKEIV
GOHGGAHTYTIGORGGLDLGGSEKPCYVVGKNIEENSIVIVRGEDHPQLVLRELTAREIN WFTPPKSGCHCSAKVRYRSPDEACTIDYSSGDEVKVRFSQPVKAVTPGQTIAFYQGDTCL GSGVIDVPMIPSEG

CPn_0439 485523 486077
No robust homolog present in Genebank/EMBL as of 11/7/98
IISSNIB'/LFVSSTLNGVFPSSLPEESADLFITNKEIVALGEKGNVFLTHSIPMHIAAIT
ILVIVALAGIAIICLGCYSQSILLIAVGIVLTILTLLCLQALVGFIKFIRQLPQDLHTTV OF IREK IRPESSLOLVTNAORKTTODTLKLYEELCDLSOKEFKLOSTLYOKRFELSHKNE

CPn 0449 486081 - 486740 CPn_0449 486081 486740

for robust, homolog present, in Gerebank/EMBL as of 11/7/98
LATTRERRMATSVARSTVPESOPLISHATEVENTLENAYHTOPHPIPAAPWETFRSKLSTKH
TLEFALTELLTEXTIBAGYAGYTONWITCGTGLGITVTT.LILALLATTEKNKOTKTKL
IDEIGGITGIGGTVORYGLMETTIK:WHILELTTGNOEKTRILNETEAKKEGIQNLEL
KITEGGITLAQKQPKRKSSQKGFMRGIKHLSKNPVILFDX

486875 4878 (8 CERCO441 ASSAUL ARVESS
CTRO7 CYPOT BETTA ASSAUL ARVESS
CTRO7 CYPOT BETTA ASSAUL ARVESS
WATFOUT CYPOT BETTA ASSAULTED FOR THE CONTROL OF THE PROPERTY C

GEYYWS LGND LS LADDH

489528

CTOUS hypothetical protein NILEGYTHMGFKNICKGOSGLYLHVIFPERILARKLKNCAKSYPRTALTIEVLVSSVLGAL KVILIFCASTYAALTLEDHRALFHAIKTKGCGHLASYAMAWLLHILTIAVIIGLVFSLVFI PPPVVFISLDILLMGVTTGVTLFG/MKNLFPPYEPPBRPHTPPPFADEYVPLISESYPD

4 087.5 489973 CT005 hypothetical protein

CT005 Mypothetical protein
VDSMSQPPINPLCQPQVPAAASPSCQPSVVKRLKTSSTGLFKRFITIPDKYPKMRYVYDT
GIIALAAIAILSILLTASCNSLMLYALAPALALGALGVTLLISDILDSPKAKKIGEAITA
IVVPIIVLAIAAGLIAGAFVASSGTMLVFANPMFVMGLITVGLYFMSLNKLTLDVFRREH
LLRMEKKTQETAEPILVTPSADDAKKIAVEKKKDLSASARMEEHEASORQDARHRIGRE
AQGSFFYSSRNPEHRRSFGSLSFKFKKPSDAASTRPASISPPFKDDFOPYHFKDLRSSF
GSGASSAFTPIMPASSRSPNFSTGTVLHPEPVYPKGGKEPSIPRVSSSSRRSPRDRQDKQ
QQQQQQDEEQKQQSKKKSGKSQSLKTPPPDGKSTANLSPSNPFSDGYDEREKRKHRKNK

CPn_0444 490266 494507
pmp_6-Polymorphic Outer Membrane Protein
KAFPQRHMKYSLPWLLTSSALVFSLHPLMAANTDLSSSDNYENGSSGSAAFTAKETSDAS
GTTYTLTSDVSITNVSAITPADKSCFTNTGGALSFVGADHSLVLQTIALTHDGAAINNTN
TPALSFSGFSSLLIDSAPATGTSGGKGAICVTNTEGGTATFTDNASVTLQKNTSEKDGAAV
SAYSIDLAKTTTAALLDQNTSTKNGGALCSTANTTVQGNSGTVTFSSNTATDKGGGIYSK
EKDSTLDANTGVVTFKSNTAKTGGAWSSDDNLALTGNTQVLFQENKTTGSAAQANNPEGC EKDSTLDANTGVVTFKSNTAKTGGAWSSDDNLALTGNTGVLFQENKTTGSAAQANNPEGC
GGAICCYLATATDKTGLAISQNQEMSFTSNTTTANGGAIYATKCTLDGNTTLTFDQNTAT
AGCGGAIYTETEDFSLKGSTGTVTFSTNTAKTGGALYSKGNSSLTGNTNLLFSGNKATGP
SNSSANQEGCGGAILAFIDSGSVSDKTGLSIANNQEVSLTSNAATVSGGAIYATKCTLTG
MGSLTFGNTAGTSGGAIYTETEDFTLTGSTGTVTFSTNTAKTGGALYSKGNNSLSGNTN
LLFSGNKATGPSNSSANQECCGGAILSFLESASYSKKGLWIEDNENVSLSGNTATVSGG
AIYATKCALHGNTTLTFDGNTAETAGGAIYTETEDFTLTGSTGTVTFSTNTAKTAGALHT
KGNTSFTKNKALVFSGNSATATATTTTDQEGCGGAILCNISSSDIATKSLTLTENESLSF Inntakrsgggiyapkcvisgsesinfdgntætsggalysknlsitangpvsftnnsgg kggalyladsgelsleaidgditfsgnrategtstpnsihlgagakitklaaapghtiyf KGGATYIADSGELSLEAIDGDTFSGNRATEGTSTPNSIHLGAGAKITKLAAAGHTIYF
VDPITMEAPASGGTIEELUINPVVKAIVPPPOPKNGPIASVPVVPVAPANPNTGTIVFSS
GKLPSQDASIPANTTILNQKINLAGGNVVLKEGATLQVYSFTQQPDSTVFMDAGTILET
TTTNNTDGSIDLKNLSVNLDALDGKRMITIAVNSTSGGLKISGDLKFHNNEGSFYDNPGL
KANLNLPFLDLSSTSGTVNLDDFNEIPSSMAAPDVGYQGSWTLVPKVGAGGKVTLVAEWQ
ALGYTPKPELRATLVPNSLWNAYVNIHSIQQEIATAMSDAPSHPGIWIGGIGNAFHQDKQ
KENAGFRLISRGYIVGGSMTTPQEYTFAVAFSQLFGKSKDYVVSDIKSQVYAGSLCAQSS
YVIEHSSLRRHVLSKUYPELPGETPLVLHGQVSYGRNHHMMTTKLANNTOGKSDMDSH
FAVEVGGSLPVDLNYRYLTSYSPYVKLQVVSVNQKGFQEVAADPRIFDASHLVNVSIPMG
LTÜĞGESAKPPSALLLTLGYAVDAYRDHPCLTSLTNGTSWSTFATNLSRQAFFAEASGH
LKEHGLGLOFFASGSCELRSSSRSYNANGGTRYSS LKLEHGLDCFASGSCELRSSSRSYNANCGTRYSF

CPL0445 494739 497579
pm57-Polymorphic Outer Membrane Protein
preflyskkclonkssyswiffssiplfssisivaaevildssnnsydgsngttptvfstt
paakattysllsuvsfonagalgiplasgoftleaegglitpognohalkfafinagssagt
väefsaadknllfndpsrlsiiscpslllsptogcalksvgnlsltgnsqiiftonfssd
ngyintknfllsgtsofasfsrnoaftgkoggvyatattitiensgiisfsonlakgs
ggalkstdncsitdnefovifdonsaweaaqoqoggaicttidktvitutiknknlsfinnt
liyegaiscikvsisaggfilfosnisgssaggoggainiasagelalsatsgdiffnn
novingststrnainiidtakvtsiraatggsiyfydpitnpgtaastdtilnladans
eifyggaivfsgeklsptekaiaanvtsiropavlargdlvlrdsvtvffkdltospgs
rilmdogttlsakeanlslnglavnlssldginkaalkteaadknislsgtialidtegs
rilmdogttlsakeanlslnglavnlssldginkaalkteaadknislsgtialidtegs
ryeninlksastypllelitaganstitlolestitlopepethygyognwolswanatss
kigsinwtrtgyipsperksnlplnslmgnfidirsinglietkssgepferelmlsggia
nferdsmytreggfpinsprekafaamvuseisgiiplespakfsylhtnohmktytyto
siikgswrdafcadlgaslpfvisvpyllkevepfvkvoyiyahoqdfyerhaegrafy
kselinveipigvfferbsksekgtydltlmyildayrnprkcotslisdanwmaygtn
largsfsvraanhfovnphmeifgofafevrssssrnyntnlgskfcf 494739 497579

CPHE0446 497602 500415
pmp_8-Polymorphic Outer Membrane Protein
LIEEHLISKKIPLHKLLISSTLVTPILLSIATYGADASLSPTDSFDGAGGSTFTPKSTAD
ANGTNYVLSGNYINDAGKGTALTGCFTETTGDLFTTGKGYSFSFNTVDAGSNAGAAAS
TTADKALTFTGFSNLSFIAAPGTTVASGKSTLSSAGALNLTDNGTILFSONVSNEANNNG
GAITTKTLSISGNTSSISTFTSNSAKKLGGAIYSSAAASISGNTGQUVFNNNKGETGGGAL
GFEASSSITONSSLFFSGNTATDAAGKGGAIYSEAAASISGNTGQUVFNNNKGLFFEGGAL
GFEASSSTTONSSLFFSGNTATDAAGKGGAIYSEAAASISGNTGUVFNNNKGLTFAENSSVTO
GGAICAHGLDLSAAGFTLFSNNRCGNTAAGKGGAIAIADSGSLSLSANGGDITFLGNTLT
STSAPTSTRNAIYLGSSAKITNLRAAQGOSIYFYDPIASNTTGASDVLTINOPDSNSPLD
YSGTIVFSGEKLSADEAKAADNFTSILKQPLALASGTLALKGNVELDVNGFTGTEGSTLL
MOPGTKLKADTEAISLTKLVVDLSALEGNKSVSIETAGANKTITLTSPLVFODSSGNFYE
SHTINQAFTOPLVVFTAATAASDIVIDALLTSPVOTPEPHYGYGGHWEATWADTSTAKSG
TMTWYTTGYNPNPERRASVVPDSLWASFTDIRTLOOIMTSQANSIYQORGLMASGTANFF
KKDKSGTNQAFRHKSYGYIVGGSAEDFSENIFSVAFCQLFGKDKDLFIVETTSHNYLASL
YLQHBAFLGGLPMPSFGSITDMLKDIPLILNAUGSYSTYKNDMTRYTSYPEAQGSWTNN
GGALELGGSLALYLPKEAPFFQJYFPFLKFQAVYSRQQNFKESGAEARAFDDGDLVNCSI
PVGIRLEKISEDEKNNFEISLAYIDDVYRKNPRSRTSLMVSGASWTSLGKNLARQAFLAS 497602 500415

CPn_6447 500541 503351
pmp_9-Polymorphic Outer Membrane Protein
PYMPPIALYMKSSLIMFLISSELDPLSLNFSSFAAVVEINLGPTNSFSGPGTYTPPAQT
TNADGTIYNLTGDVSITNAGSPTALTASCFKETTONLSFQGHGYOFLLONIDAGANCTFT
NTAAIFLLDFSGFSYLDLIQTTNATTGTGAIKSTGAGSIGCNYSGYFGQNFSNDNGGALQ
COSTGLIGHNNLTFAKNKATOKGGALYSTXGITIMITLNSAGSFSENTAANKGAIYTEAS NTAAII LLOFEGEGYLÖLLOTTNATTOTGA I KSTGAGS I QONYS Y FEQNESNINGGALO GOS I GLOLIPRILIFAKNATOKGALYTGAT I INTILISASESENTAANIGGALYTEAS OF I COLIVA LEE I NIEUVIATOKATAJA I YGOSTGARPEPUTLUDOMELIRE I GITTA I TOSAGA I YTEADA I ALOCOGLISA GETURA LEE INTILISASESENTAANIGGA I YTEADA I ALOCOGLISA GETURA I GETURA I GETURA I GETURA I DE ALOCOGLISA GETURA I YGOSTGARA I ALOCOGLISA GETURA I GETURA

SYNADLANKFOF GIFYAFSPNLEVTCHLSME

CPn_0449 594876 503698
*yxjG_Bs_2 Hypotherical Protein
FIOPSRREIHEWKCILLSSLEMEMMSFFQOPEOCHFDVVGSFLRPESLTRARSDFEEGR
IVVEOMRVVEDAAIPNLIKKOTEACLIFFTDGEFRRYSWDFDFMWGFWGVDRRRDSNDPE
IGVYLKDKISVSKHPFIHFEFVKTFEKGNAKAKGTIPSPSOFFHEMIFAPNLKNTRKFY PUNDELIDDIVETY/POVIQUELYAACCRILQLDCAWCELLDIRAFSWY/VDSHDELQELI PON MINIMAMELAETLEFULHA PODYOATER PEANYO LEEDILEARDVDCHYNA LEEDIA MAEDIACOO (ERFOTELICIERINEATER PEANYO PEYRAA TI LEGERLDLIFC CGFASCEGDHRMTEEZOWKK!AFVKE!AKE!WG

CPn_0449 507231 505330
pmp_10-PMP_10 (Frame-shift with 0451)
EAYTGFRGGGISFSNNIVOGTTAGNGGAISILAAGECSLSAEAGDITFNGNAIVATTPO
TTKRNSIDIGSTXY.TNLTAHISGRISIFYDPITANTAADSTDTLINKADAGNSTDYSGS
TVFSGEKLSEDEAK/ADNLTSTLKOPVYDTAGNIVLVLKGYTLDTKGFTQTAGSSVIMDAG
TTLKASTEEVTIJGLSIPVDSLGEGKKVVIAASAASKNVALSGPILLLDNOGNAYENHDL
GKTQDFSFVOLSALGTATTTDVPAVPTVATPTHYGYQTWGHTWVDDTASTPKTKTATLA
WINTGYLENPFROGPLVENSLWGSFSDIQAIGGVIERSALTLCSDRGFWAAGVANFLOK
KKGEKRKYRHKSGG/AIGGAAQTCSENLISFAFCOLFGSDKDFLVAKNHTDTVAGAFYIQ
HITECSGFIGCLLEKLPGSWSHKPLVLEGOLAYSHVSDLKTKYTAYPEVKGSWGNNAFN
MMLGASSHSYPEYLHCFDTYAPYIKLNLTYIRQDSFSEKGTEGRSFDDSHFNLSLPGV
KFEKFSDCNDFSYDLTLSYVPDLIRNDPKCTTALVISGASWETYANNLARQALQVRAGSH
VAFSPMFEVLOGFVFEVRGSSRIYNVDLGGKFQF

CPn_0490 508121 507180
pmp_10-Polymorphic Outer Membrane Protein
scpmks-OFSWLVLSSTLACFTSCSTVFAATAENIGPSDSFDGSTNTGTYTPKNTTTGIDY
TLTGDTTLONLGDSAALTKGCFSDTTESLSFAGKCYSLSFLNIKSSAEGAALSVTTDKNL SLTG/SSLTFLAAPSSVITTPSGKGAVKCGGDLTFDNNGTILFKODYCEENGGAISTKNL SLKNSTGSISFEGNKSSATGKKGGAICATGTVDITNNTAPTLFSNNIAEAAGGAINSTGN CTI/GNTSLVFSENSVTATAGNGGALSGDADVTISGNQSVTFSGNQAVANGGAIYAKKLT

CPL_0451 508158 511058
pdp_10-PMP_10 (Frame-shift with 0451)
RTORVKIKILDSCFVIFNLIYLFCFYIDANSSLKNKSITMKTSIPWVLVSSVLAFSCHLQ
SLANEELLSPDDSFN3NIDSGTFTPKTSATTYSLTGDVFFYEPGRGTPLSDSCFROTTDN \$LANEELLSPDDSFNGNIDSGTFTPKTSATTYSLTGDVFFYEPGKGTPLSDSCFKQTTDN
LTFLGNGHSLTFGFIDAGTHAGAAASTTANKNLTFSGFSLLSFDSSPSTTVTTGGGTLSS
AGGVNLENIRKLVVAGNFSTADGGAIKGASFLLTGTSGDAIFSNNSSSTKKGGAIATTAGG
RIANNTGYVRFLSNIASTSGGAIHAKKLALSSGGFTEFLRNNVSSATFKGGAIATTAGA
RIANNTGYVRFLSNIASTSGGAIHAKKLALSSGGFTEFLRNNVSSATFKGGAIATIGA
ELISINAKTLIFASNVAETSGGAIHAKKLALSSGGFTEFLRNNVSSATFKGGAISTDASG
ELISINAEGNITFYRNTLTTTGSTDTPKRNAINIGSNGKFTELRAAKNTHIFFYDPISIDASG
GTSSDVLKINNGSAGALNPYQGTILFSGETLTADELKVADNLKSSFTQPVSLSGGKLLLQ
KGVTLESTSFSQEAGSLLGMDSGTTLSTTAGSITITNIGINVDSLGLKQPVSLTAKGASN
KVIVSGKLULIDIEGNITYESHMFSHDQLFSLLKITVDADVDTNVNISSLIPVPAEDPNSE
YGFQGQWNVNTTDTATNTKEATATWTKTGFVPSPERKSALVCNTLWGVFTDIRSLQQU
EIGATGMEHKQGFWSSMTNFLHKTGDENRKGFRITSGGYVIGGSAHTFKDDLFFTAFCH
LFARDKDGFIAHNNSRTYGGTLFFKHSHTLQPQNYLRLGRAKFSESAIKFPREIPLALD
VQVSFSHSDNRMETHYTSLPESEGSWSNECIAGGIGLDLFVLSNPHDLFKTFIPOMKVE
MYYVSQNSFFESSSDGRGFSIGRLINLSIPVGAKFVQGDIGDSYTYDLSGFFVSDVYRNN
POSTATLVMSPDSWKIRGGNLSRQAFLLRGSNNYVYNSNCELFGHYAMELRGSSRNYNVD
VGTKLRF VCTKLRF

CPn_0452 511304 512860
pmp_12-Polymorphic Outer Membrane Protein (truncated)
pme_12-Polymorphic Outer Membrane Protein (truncated)
pneETMT1LRNFLTCSALFLALPAAAQVVYLHESDGYNGAINNKSLEPKITCYPEGTSYI
FLDDVRISNVKHOEDAGVFINRSONLFFMGNRCNFTFHNLMTEGFGCAISINRVGDTTLT
LSNFSYLAFTSAPLLPQGQAIYSLGSVMIENSEEVTFCGNYSSWSGAAIYTPYLLGSKA
SRPSVNLSGNRYLVFRDNVSGGYGGAISTHNLTLTTRGPSCFENNHAYHDVNSNGGAIAI
APGGSISISVKSGGLIFKGNTASQDGNTIHNSIHLQSGAQFKNLRAVSESGVYFYDPISH
SESKKITDLVINAPEGKETYEDTISFSGLCLDDHEVCAENLTSTILQDVTLAGGTLSLSD
GVTLQLHSFKQEASTLITMSPGTTLLCSGDARVQNLHILIEDTDNFVPVRIRAEDKDALV
SLEKLKVAFEAYWSVYFFPQFKEAFTIFLLELLGFSFDSLLLGETTLERTQVTTENDAVR
GFWSLSWEEYPPSLDKDRRITPTKKTVFLTWNPEITSTP

CPI_0453 513156 516152

pmp_13 -Polymorphic Outer Membrane Protein
NCVLLYLFFYSLSLICRIIWFHLYVOMKTSIRKFLISTTLAPCFASTAFTVEVIMPSENF
DCSSCRIFPYTTLSDPRGTILIFSGDLYIANLDNAISRTSSSCFSNRAGALOILGKGGVF
SFLNIRSSADGAAISVITONPELCPLSFSGFSOMIFDNCESLTSDTSASNVIPHASAIY
ATTPMLFTNNDSILFQYNRSAGFGAAIRGTSITIENTKKSLLFNGNGSISNGGALTGSAA
INLINNSAPVIFSTNATGIYGGAIYLTGGSMLTSGNLSGVLFVNNSSRSGGAIYANGNVT
FSNNSDLTPQNNTASPQNSLPAPTPPPTPPAVTPLLGYGGAIFCTPPATPPPTPGVSLTIS
GENSVTFLENIASEQGALYCKKISIDSNKSTIFLGNTAGKGGAIAIPESGELSLSANQG
DILFNKNLSITSGTFTRNSIHFGKDAKFATLGATGGYTLYFYDPITSDDLSAASAATVV
NPKASADGAYSGTIVFSGETLTATEAATPANATSTLNOKLELGGGTLALRNGATLNHV
FTQDEKSVVIMDAGTTLATTNGANNTDGAITLNKLVINLDSLLGGTKAAVVNQSTNGALT
ISGTLGLVKNSGOL-JONIGMFNKBLOQVPILELKATSNTVTTTDFSLGTNGVQOSPEVG
GTWEFT IDTTTHTLTGNWKKTGYLPHPERLAPLIPNSLWANVIDLRAVSQASAADGEDVP
GKOLSITGITTMFFHANHTGDARSYRHMGGGYLINTYTRITPDBALSGFGQLFTKSKDYL
VGHGHSNYYFATV-SNITKSLFGSSFFSGGTSRVTYSRSNEKVKTSTYTKLPRGRGSWN
NCWLGELEGNLPITLSSRILNLKGIIPFVKAEVAYATHGGIOENTPEGRIFGHGHLLNVA
VPVGVPGRCKNSGNREDPYTIIVAYAPDVYRHNPDCDTTLPINGATWTSIGNNLTRSTLLV
QASSHTSVNDVLEIFGHCGCDIRRTSRQYTLDIGSKLRF

CPn_0454 516179 519115
pmp_14-Polymorphic Outer Membrane Protein
GMPLSFKSSSFGLLAGLGSASCAFAETRLOCHFYPPITHOGEEILLTSDFYGENFLGASF
SSSFINSSSNLGLLGKGLSSFFTSCOAPPINSHYALLSAAETLTFKNESSINFTCNOCTGL
GGLTYGKDIVPOSISULIFTTNRWAYSPASYTTSATPAITTYTTGASALGHTDGLTVEN
SOCIFFENIANFGSAIGSSFTAVVKFINNTATMSFSHNFTSSCGVIYVGSGLLFENNS
GGLIFTANGGVINS GAVTPSSGTYALGSGGAIGIPTGTFELKNNOGKCTESYNTTPNDAG
ALVAFTUNIVGBNESSILD DESMFAABANGAAIGEVININTGGGGLIFFSSNRAFKGCAIGEKG CALL LETANDOME: RGYTESSOTYALGSOGALC LETGTEELKNNOGK-TESYM-TEPHDAG
A LYABTUN VONCALLEDENTAANDOMA LOAVUN LOPGE LEFERNRAEKOAA LE LIGE
SVODPAKOTUTELT LAGRID LAFONMILITER EL RINA LTVEAKKE VISLSAOV KURLIVEY
DE LTHIGLETTOISSIND LETANDOMENE VVETOR OLGSTEELLE PANTITELLETVK LAGGE
LK LTHIAVVITZIGEAT CHOQUETUR STELLEGATPTOAPAAVOET LIKLAEDDESELKRO
EVOSATINA TENNITOIT GALTLEBANTOLLIKAN ZOOSPIVAL DE TAVEKGATVITK DEPLOGE
LATPOHYGYOKWSYTWORDLIL PAPLIK JEFF ZIPGEGATTLYAVWNOTELVRITET LLDEE
RYGELYONOLWESELRIOAFSOLLODVILLEDHE GLOGETTAKALGAYVEHT PROMIEGEOOR YCHYOAALHMEYTDHTTLGUSFIQUYGKTNANPYDSR LEISWKAAYGYJKHILNTTYLRPDKAPKSQOOWHNISY WDLIGG ISABELGCWGKFTETGDLQRSFSRGKGYNVS KLAYKPOLYRWIPHIN LYTVVSNQESTS ISGANLRRHGLFVQIHDVVDLTEDTQAFLNYTF LGKNGFTNHRVJTGLKSTF

CPn_0455 520363 519458
No robust homolog present in Genebank/EMBL as of 11/7/98
ROBUSTO REPLANDED PROPERTY AND REPLANDED LEELS PROVIDED TO REPLANDED PROPERTY AND REPLANDED PROPERTY AND REPLANDED PROPERTY ASCRIVATED BEHAVIOLE REPLANDED PROPERTY SERVING ALDAVI KOSVWRFRNDLFAYEREALEASVTDALVSYVSNLDMIPYTSSQGIVIEDSSIV
RTSQEHTLIVNCAAFDKLASQIEFLCPSDVLPISGKDPLISDDEDEELNPKVSSAADSKD

CPn_0456 521568 520327
No robust homolog present in Genebank/EMBL as of 11/7/98
IPCTFESKRKFLMTHCLHGWFSVVRHHFVQAFNFSRPLYSRITHFALGVIKAIPIVGHLV
MGVWLISHCFERGVSHPGFPSDIAPILKVEKIAGROHISRIENQLKSLRKTIEVEDLDK
VHGQYQENPYADMASSEVLKLDKGVHVSELGKAFSRVRNRITRSYSYAPTPQLDSIAIVG
IDLVSPEEQENLVRLANEVIQLYPKSKTTLYLLIDFNKEWVGDISSDKEKQLRSLGLHSE
VQCLSVLEPQGASGEDTHKFDLMVGCYGKDSYLREGKILQQALGTSIGTIVFWNVMHTLP
SRYRSRLSLPINTEKDKTELYKEISRTHHQLHTLGMGLGAQDSGLLLDRGRLHAPLSQGS
HCHSYLADLTHEELKILLFSAFVDAKNISKKELREVSLNFANDTSVECGCAFYF

CPn_0457 523886 522120

No robust homolog present in Genebank/EMBL as of 11/7/98

VFLPSRVMASCLSAWFSIVREHFYRAFDFSLFPCARITEFVLGVIKGIFVVGHIIVGIEW

LVSRYLESFVTKFFFVSDVVSLLKTEKVAGROHIARVVETLKRQRVAVAPEDEDKVHGKI

FVHPFGGIQPVEVLTLYPEVQDATIGLAFSKIRNRVRQAYLQAPRPKLQKIYIIGNDMNP

FEVDDFLHLARLCNETQRLYPDATISLYLTASGGRNAMDKKNRKLLSDCELNPKIACLDF

NQGDVVKQATCDCWMVYHGENDQGTLIXQIGELEKSGEFTPHIHVGQKPLSGSLWDFSPF

SSLEMKGDKEKALEYSELEKEQLYSRLVYVGERSSVLSLGFGDSRSGILMDFKRVHAPLS

EGHYCHSYLADLENPGLQKTILAAFLNPKELSSTILQPISLNLILNSKTYLRQHFGFFER

MSRSDRNVVVVCDSWAGTDWKEEPSFOHFIMELERGYSHFNIFAFRSNGWCYERRIL

NESSQEKAFTMIFCEDSVSQGDIRCLHLASEGMLCGKECYAVDVTTSCCANFMMEEVLTL

ERESNLWNRKHGLWKREVYKQKQEAALDQDESEIYVCNQLTAQQNFACS

CPn_0458 526344 524236
No robust homolog present in Genebank/EMBL as of 11/7/98
YFECCYLKLFYSNFIFFVVMPIPYISSWISTVRQMFVKAFDFSRFCSRVTNFALGVIKA
IP#VGHIVMGMEWLVSSCVAGIITRSSFTSDVVQIVKTEKALGRDHISRVAEILQRERGT
IP#NGHVMGKFPVCPFGRLKSEETLKLKPGEREGTLDTVFSPIRTRVTRAYLQAPRPE
IRTISIVGSKLKTPODFSOFVSLANETORLHPEALVCLYLTGLNRESGMCDTTTAEKKQY
LHRSGLDSRICCKDSKEDDAGSPEDREJKIGTYSREQOHNIDGOVIQOCLGKSADPIWH
VTEDTKDFYYPPNFTSYSHTRQSTDPTSPPRLPESEGDKDSLYGQLSRSYHHEYMLGLG
LKPEDDAGLLMDPDRIVAPLSGGGHYCHSYLADIENEDLRTLVLSFFLDFGNLSSEDLRPVA
FNIARLPLELDSLFFRLVAGGODGSNIVTLAHGTRPEDLDPDSNILITRRLONGGYSYL
NIFSYKSRKMIVKERQFFGDRSEGKSFTLILFEDPISAADFRCLQLAAEGMVAKDLFSVA
DICASGCSCIQFSEMOSPQAIEYRGWEARVEDEAGEERREPVIYSQDQLSSMLTTQONFV
FSÜDAVKOAINWFRSKGLLTHERKALGEFLTATFSYLGSGERNENMGKRTTEEHEVVI
SEEELDRMVQVLPAEVPADSGNDPTRPVPNPDSNPDSSQNEGS

CETT0459 527062 526619
NOTFOBUST homolog present in Genebank/EMBL as of 11/7/98
STKIQHHPGLRNWRTSTNKLREEGSVSFREYFRAYMCDKIVAQKNFLFTLDAVIKQAGWR
SQEKLNLFYVESQALGREIKVSLEEYIQSMVGILGSQRTKKSFKFSVDFTPLEQALQERC
SGEDDEDATATSTATGATASPTDMHEDE

CHE-0460 527840 526992

No robust homolog present in Genebank/EMBL as of 11/7/98

VEHILNFALEETPSISVQYQEQEKLSPCDHSPEIGKKKRWNKLESFSTYCSLFMSVKDM

YKINLGIQNSLSGWLLDPYRVCAPLSSPYSCPSYLLDLONKELRRSLLSTFLDFKNLTSE

TFKEVSINFONSSFQQRWSEFLSRVLHDEKEKHVAVVCNDAKLLEEGLSPEALSLLEEDL

REGGYSYLNILSVSPEXOKNOCEGLIRRDLQGRSFTVMITDLPIGSEDIRSLQLASFRI

LVSSSLDAADACASGCKVLVYENPNASWAQELENFYKQVERRR

CPHE0461 528647 527844
No robust homolog present in Genebank/EMBL as of 11/7/8
ISIVACPSISSWFTVVRQHFVNAFDFTHPVCSRITNFALGIIKAIPVLGHIVMYIEWLIS
WIPRHTVRHGMFTSDVSSAIKVEQTRGHNCLAPLEAYLSSLRVPISQEDLGKVHGRTPED
PFVDITPTEIVQLLPDEELSTVDEALOGVRSRITYAYRSVEKPMIQDLALVGTGLRDSAD
LINFVRLANGVQNHYPHTKVKLYLAKNLADVWDCEISEEKGQLRALGLDPWIESISLTS
AGLPSVPEVATVDFMITCYGKDQEVQDP

CPn_0462 531124 529037
No robust homolog present in Genebank/EMBL as of /1177/98
LIFYLFLNLYIACVRFHFQCWFDPMACYISIWISTVKQHFIRAFDFTRFLGSRITNFALG
VIKAIPILGCVVIGVSWLVSTCSARRFGKPAFTSDVASIVKIEKTRGKNPLAWVEQVLRQ
LRVRLPEGDLGKHGKVSRDYVCDRTPQENLMVPHQVLGELGRAFYGIRNRVTKAYGRV
TPLEVPCLTLVGFDILDPEDQVNFVRLANGIQTQYPQTQIKLYLLISIOKIMNQCDGTISQ
EKEQQLRSLGLDAKIKCVSAPALLLQKYLQSENLPSCDLLINYYGKQQSVRDVDSIKSLL
NLSSEHIPAISVTYRPDDFFYSYYFFPGSQGTAPDQRIPMSEQFHLQTYTTLSNPRCDR
YAVHLGMEDPASGVFLDPLRVSAPLSGEYSCPSYLLDLKSEELFFLLSAFIDPNNSGQG
NPRPMSINFGNSPLGQRWSEFLSRVLHDETEKHVAVVCNNPQL/KKSFPSHSLSLLENEL
EEGGYSYLNIVSVSQERTCVKERRILSSDPSGRSFTVILTDL/EGSSDIRNLQLASDRIL
VGSALDAADACASECKILEYEDPEDEDWAQOYASFYRNIDRAGUQRQGIPGEPLGVSAST
RVVLEKDIVFNLNAVIQ\AMWKFKKRDLFAVESQALGDDMRALEGYIGSSLLVEGTIQP
OVACNVNVSPATLDEAVCAACDSAQDAPSEENNTDD

CPH_0463 532480 531191

TO TOBUST HOMOTOR PRESENT IN GENEBANK/EMBL AS OF 11/7/98

LEGGYPERTROLLGTPRICHTPRIVALETVOLE LEDETCHAFVOSHMIKOTVODAKELYTFLSR

RIGHTYGELE LAWSTHERE LEDET FORKMICAR PLESTONAFVOSHWIKOTVODAKELYTFLSR

RIGHTYGELE LAWSTHINGDSFORLOGKDELSMULRDETY KNOVUVYEKOVLOLE PATOVCK LVEE

LAWSTOTYLIN FERDICIDSSFORLER RELECTORY TVY CALYLIDTDMRSLOLASERIM

VORSEBLINDAYAAMGELLE (DHTWAR POTTSERIADE ADAVOVSACEPHSREERKLITOANOG

LLEGGELGLESKTEWRIGHLAFOROTVTTRIFT LPMLDAALKGANWTHKHESLIDKECEALD

LAMSTELLESTVSTLESVTNISHEKTOKOPFLOKE LIADCOPLKEALFPGSDEDV POTSEDPS

BAHEGOLEDS

CPn_0454 533366
No robust homotod planet in Genebink/EMBL is of 11/77/99
SLETRGRFTEICLOLLFEDIGEKFLOLFEBOTALNLER FEADLRNRYTTEYSRARQPDL
HRIAIVYIGVLOSESSKILERLISYMSC/YSESOMYLRFFMCKNYNGSAVLSKLHVENLH
IRCGFFSEDAVPESEPFDLSIYVHTDRSCPLPTKKRSSSWELCTVELPESIYPQSEFLLM
ROBMY S

CPn_0465 533278 512871

The testern boundary process of the testern boundary process of the testern boundary process of the testern boundary process of the testern boundary process of the testern boundary process of the testern boundary process of the tester boundary process of

CPn_0466 533/18 536537

pmp_15-Polymorphic Oyter Membrane Protein
TSMRFFCFGMLLPFTFVLAMSGLOLPLETYTILSPEYQAAPQVGFTHNONQDLAIVGNHN
DFILDYKYRSNGGALTKKILISENIGNVFFEKNVCPNSGGAIYAAQNCTISKNONYAF
TTNLVSDNPTATAGSLLGGALFAINCSITNNLCGCTFVDNLALINKGGALYTETNLSIKDN
KGPIIIKONRALNSDSLGGIYSGNSLNIECNSGAIQITSNSSGSGGGIFSTOTLTISSN
KKLIEISENSAFANNYGNFNPGGGGLTTTFCTILINNEGVLFNNNQSQSNGGAHAKSI
IIKENGPVYFLNNTATKGGALLNLSAGSGNGSFILSADNGDIIFNNNTASKHALNPPYRN
AIHSTPNNNLQIGARPGYRVLFYDPIEHELPSSFPILFNFETGHTGTVLFSGEHVNONFT
DEMNFFSYLRNTSELRGCVLAVEGAGLACYKFFQRGGTLLLGQGAVITTAGTIPTPSST
PTTVGSTITLNHIAIDLPSILSFQAQAPKIHIYPPKTGSTYTEDSNPPIITISGTLTLRNS
NNEDPYDBJULSHBLEKVPLLYIVDVAAQKINSSOLDLSTINSGEHYGYGGIMSTYWVET
TTIINPTSLLGAYTKHKLLYANMSPLGYRPHPERRGEFITNALWOSAYTALAGLHSLSSW
DEEKGHAASLQGIGLLVHQKNAGFKGFRSHMTGYSATTEATSGSPPSSUGFGFFSKA
KEHESQNSTSSHHYSGKCIENTLFKEWIRLSVSLAYMFTSEHTHTMYGGLLEGNSQGSF
HNHTLAGALSVFLPQPHGESLQIYPFITALAIRGNLAAFGESGDHAREFSLHRPLTDVS
LPVGIRASWKMHRVPLLWLTEISYRSTLYRODPELHSKLLISQGTWTTQATPVTYNALG
IKVKNTMOVFRVTLSLDYSADISSSTLSHYLNVASRMF

CPT_0467
536528 539434
pmp_16-Polymorphic Outer Membrane Protein
NEILTISDONRKIKEPLVSKTPPKFLFYLGNFTACMFGMTPAVYSLOTDSLEKFALERDE
EFRISSPELLDSLSTLIGFSPITTFFGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCN
NFLLSWCDHAFFSKNLAIGTGGATACQGACTITKNRGPLIFFSNRGLNNASTGGETRGG
ALACMOPFISONOGFFYFVNSVNNWGGALSTNGHCRIQSNRAPLLFFNNTAPSGGGAL
RSDYTTISDNTRPIYFKNNCGNNGGAIGTSTVYAKINNGSVJFFNNTALSGSINSGNGS
GGAYTYTNLSIDNNPGTILFNNYCIRLGGAICTOFLTIKNSGHVYFTNNGNNGGALML
LODSTCLLFAEQCNIAFONNEVFLTTFGRYNAIHCTPNSNLQLGANKGYTTAFFDPIEHQ
HYTNPLIFNPNANHGOTILFSSAVIPEASDYENFISSSKNTSELRNGVYSIEDRAGWQ
FYKFTQKGGILKLGHAASIATTANSETPSTSVGSQVIINNLAINLSPSILAKGKAPTLWIR
PLOSSAPFTEDNPTITLSGPLTLLINEENRDPYDSIDLSEPLONIHLLSLSDVTARHINT
DNFHPESLNATEHYGYGGIWSPYWETITTTNNASIETANLLYRALYANMTPLGYKVNPE
YQGDLATTPLWOSFHTMFSLLRSYNRTGDSDIERFFLEIQGIADGLFYHQNSIFGARGFR
IQSTGYSLQASSETSLHQKISLGFAQFFTRTKEIGSSNNVSAHNTVSSLYVELPMFOEAF
ATSTVLAYGYGDHHLMSLMPSHQCABETCVSHTLAAAIGCSFPWQKSYLHLSPFVQAI
AXRSHOTAFEEIGDNPRKFVSQKFFYNLTLPLGIGGKWQSKFHVPTEMTLELSVQPVLYQ
ONPOIGVTLLASGGSWDILGHNYVRNALGYKVHNQTALFRSLDLFLDYQGSVSSSTSTHH
LOAGSTLKF

∠Pn_0468 539608 540432
pmp_17-Polymorphic Outer Membrane Protein
iyKLLDNKLMIFYDKLYPHIKVMMFMRFICLSILSTALCCSLSGNEVPNLASCOMSRKDI
SAFHTSPSRLMVTPEPLVSSFRPSNLLNGFGHDITQDITITGNSINSVIDYNYHYEDGG
ILACKNLFISENKGNLSFERNSSHSSGGALYSVRECWISKNQNYSFISNASLATTTTSG
FGGGIHALDSYITNNLGEGQFLDNVSKNRGGAIYVSLSITDNLGPIVIKKNQTLEDSS
FGGGIFCRAVNIERNYQNIQINDNNSSGGVVYFLP

CPn_0469 540399 541460
pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0469)
CFRTRGGIFSALGVIISSNKEIIEISNHSASSINTASGKLYPGGGGIMCTSLVIENNPKG
LIFNNKTAALSGGAIHTRSFIFQNNGPTAFINNSATSGGALINLSGIGSTPONFFLSADY
GDILFNNNTITSSSPOPGYRNALYAAPGINLKLGARQGYKILFYDPIDHDQTTTDPIVFN
YEPHHLGTVLFSGINVDSNATNPLNFLSKFSNSSRLERGVLAIEDRAAISCKTLSQTGGI
LRIGNAALIRTKGPGSSINFNAIAINLPSILQSEASAPKFWIYPTLTGSTYSEDTSSTIT
LSGPLTFLNDENENPYDSLDLSEPRKDIPPELPRCDCKKNRYFESHCRSHELR

CPn_0470 541357 542532

pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0470)

ISLNLERISPLLYLLDVTAKKIDTSNLIVEAMNLDEHYGYQGIWSPYWMETTTTTSSTVP EQTNTNHRQLYVDWTPVGYRPNEEHGEFIANTLWGSAYNALLGIRILPPONLKEHDLEA SLOGLGLLINQHNREGRKGFRNHTTGYAATTSAKTAARHSFSLGFAQMFSKTRERQSPST TSSHNYFAGLRFDSLLFROFISTGLSLGYSYGDHHMLCHYTEILKGSSKAFFNNHTLVAS LDCTFLPARITRTLELQPFISAIALRCSQASFQETGDHIRKFHPKHPLTDLSSPIGFRSE WKTSHHIPMLWTTEISYVFLTYRKNPEMFTTLLISNGTWTTQATPVSYNSVAAKIKNTSQ LFSRVTLSLDYSAQVSSSTVGQYLKAESHCTF

CPI_0471 542561 545401
pmp_18-Polymorphic Quer Membrane Protein
TVONNRSLSKSSFFVGALILLSKTTILLMATPLSDVFDNQANQLTTLFPLIDTLTNMTPYS
HRATLFGVRDDTNQDIVLDHQNSIESWFENFSQDQGALSCKSLAITNTKNQILFLNSFAI
KRACAMYVNGNFDLSENHGSIIFSGNLSFPNASNFADTCTGGAVLCSKNVTISKNQGTAY
FINNKAKSSGGAICAAIINIKDNTGPCLFFNNAAGGTAGGALFANACRIENNSQPIYFLN
NQSGLGGAIRVHQECIITKNTGSVIFNNFAMEADISANHSSGGAIYCISCSIKDNGDIYFV
NNRHFKDTFSNHVSVNCTRNVSLTVGASQQGISATFYDPILQRYTIONSIQKFNPNPEHLG
AFDNMTAAROGGAICTOSLTIQDSGPVYFTNNGTMGGAIMLRQDTACTLFADQGDIIFY
NNRHFKDTFSNHVSVNCTRNVSLTVGASQQGISATFYDPILQRYTIONSIQKFNPNPEHLG
TILESSTY IPDT:TTSRDDFISHERNHIGLYNGTLALEDRAEWKVVKFDQFGCTLRLSSRA
VFSTTDDERQSSSRVSVININNLAINLGSILLSTRAUPKLWIR PTGGSAPYGEDNNFIINL
SOFFISHLDDENLDFVTDALLAGFIALDVTLKILLDTAKHILTNDNFY PEGLENTTOHKYQG
WGDYWIETCITTSPSSEDTVNTLHRQLYGDWFPPCYKVNPENKGDIALSFROGFHNLF
ATLRYQTQCAGIATALSEATRLFYHIONGNNDAKGFHREATGYGATTSDNTAGNHIKAACYGGK
IQTECKKYSTTLGAALGGSLSLOWEGRHAHFPFFIOAIAVRSNUTAFQEGGDKARRESVH
KECYNELTVELIGAALGGSLSLOWEGRHAHFPFFIOAIAVRSNUTAFQEGGDKARRESVH
KECYNELTVELIGAALGGSLSLOWEGRHAHFPFFIOAIAVRSNUTAFQEGGDKARRESVH
KECYNELTVELIGAALGGSLSLOWEGRHAHFPFFIOAIAVRSNUTAFQEGGDKARRESVH
KECYNELTVELSIGENSWEKKELLFYLWHIELAYQFOLYGONFOLWEGGDETTTHTLANGTTFKF

CFn_0472 547914 545581

CPn_0473 549602 548070

NO TODUST HOMOLOG PRESENT IN GENEDANK/EMBL AS OF 11/7/98
GSIMAVCGVCGSRSPSPIPPNRRNSEDGKVSFKDNLGEHTVSSSDSSLASGGPTIEERKA
QLGGTDKIPLPSVKEPGDSGTSGRSGVLGRIKKGVKKFVFQARPEVSSPRLPSHVQH
GQRLPGLEGFRDRIQKRSENPEADLGKMKRSYSDGDLDRVCHDSNEDSTEDSRSEGGEPS
SKSSFLSGVRGAVSKVHGALGDIKGKFQRSASEDDLTTQEDSAGDTVKERRSEBABAS
SKSSSFLSGVRGATSTVQCALDDAKEKVSAFGEQAAGAIRSAFGNIRTRFQRSSSEGDLS
VKNKAAKHRKALENLEKVAPEQVSFEVASRVQSLLARMEQLTHQEPTVEDLITFVESN
VGSDSVEYASIVPQDGSQAPAETAEAPETGGVEGSAAQGAMKALRDFVVSIFQAVASFFR
AIASRLSSARRESAVDDLASESNTGMFVEQEGVSNPSAAPSLSFAEEIARRAAEMSNRNA
OSLFKLFSGNVTDPV1COGLGLARSFAPEGO QSLEKLESGNVTDPVIQQGLGLARSFAPEGQ

549807 CPn_0474 551600 549807
CT365 hypothetical protein
LKIIISISIMSTSPISNDRYLSLSNATEKTSLLANSRSLSPVPNSLVPSNPEDTGLRKS
LKIIISISIMSTSPISNDRYLSLSNATEKTSLLANSRSLSPVPNSLVPSNPEDTGLRKS
IFTHSVTLFAGLVVLLVAVSVVVVALTVLAPGVPQAILLGIAISGVGIGGFSIKKSLVYM
VRDWSPRMQESSRIKSALAVGTGFTVMGLVMKVGNFVPGGYGGLVGSLGSSAYSRGSQ
TTLASFSHYIYTKFFRSEKVAKGEKLTEAETIKEAKKLHYITLSIATIGVGLAVLGILLA
IAGTVLLGGAPATIAIILAPPLISIGLITTVLQTILHSSIGKWRAFLLTQEKKDLFVDTSL
LKILASIAAFIVTGFGGLTVMQVLLVASVGSAVASVTLPMVSSGFSYVAYQLKARLNISKL
RWKEAKNKKRVROFLIESGVIASDREFNOMMETVYKKGJGKTDAAIREEVRRFEKGGEVN
SALVGGILLGVGTGIHLLALVPAFAPIVPGILALGGSTLGIAGSILMRKFVNWLYDELVK
LYERRRNRRELLYGPESKMRSIATDLVVEALAASHDHLFDLDGPVDFIDVDVDIDGAA 551600

CPn_0475
553850
551685
glgB_Glucan Branching Enzyme
PSM*DKLIHPMDLDLLVSGRQKDPHKLLGILASEDSSDHIVIFRPGAHTVAIELLGELHH
AVA*PRSGLFFLSVPKGIGHGDYRVYHONGLLAHDPYAFPPLWGEIDSFLFHRGTHYRIYE
RMGGIPMEVOGISGVLFVLWAPHAQRVSVVGDFNFWHGLVNPLRKISDOGIWELFVPGLG
EGFEYKWEIVTOSGNVIVXTDPYGKSFDPPPOGTARVADSESYSWSDHRWMERRSKQSEG
PVTTYEVHLGSWOWDEGRPLSYSEMAHRLASYCKEMHTHVELLPITEHPLNESWGYOVT
GYARTSRYGTLOFFGVFDVTHKRHGIILDWYGHFPVDAFALASFGEPLYEYTGHS
OALHPHMNFFTFDYSRHEVTNFLLGSALFWLDKMHIDGLRVDAVASMLYRDYGREDGWT
PNA*GGKENLESIEFLKHLNSVIHKEFSGVLTFAEESTAFFGVTKDVDQGLGFDYKWNL
GWMIDTFHYPMKDPMYKHYMKDLTSLMYAPGESFILDLSHDEVVHGKGSLVMKLBGDT
WTRFAQWRVLLSYOICLPGKKLLFMGGEFGOYGEWSPDRPLDWELLHHHYMKTLRNCVSA
LNA*EN HOPYLWMGESSQECFHWJPDHIENNVIAYYRFAGSNRSSALLCVHHFSASTFP
SYVERCEGVKHCELLLMTDDESFGGSGKGNRAPVVCQDQGVAWGLDIELPPLATVIYLVT 553850 551685 VERCEGVKHCELLLNTDDESFGGSGKGNRAPVVCQDQGVAWGLDIELPPLATVIYLV

554877 553858 CPB_U4/6 5548// 553886
CT865 hypothetical protein
GRGRRADWGDCMIDIMOHFKPYTMVPGQKLPIPGSLLYAQVFPTLWRLFSSKHEILMEQT
LQVQGPLKRFAVFQDLHRGGLAVTSERYKYYLLPSGECTQSIKGKLPSAAQAGPLISLGV
HKHADWGKVRCRRDLKEILPLWFRFAAMAPKGSYRDLETTAIGSLVKTAHORVLHÆTTE
LAPALLSIALAGFSECFLERSYDEEFQGILPQDGDPEGGVPFELLSYSFGMIQDIFLRHQ
GQEVEILPALPPEPPCGRIHVALPHLGTLSIVWTKKTIRQVELHAEYSGEVFKKFCSSL CSARLREWSERRLSGSKRLSLGETLEIKAGTTYLWDCFHK

CPn_0477
•yqev_Bs 556112 554844 CPD_E9477

556112 554844

*VQEV_BS Hypothetical Protein
RYMTVAEVKGTFKLVCLGCRVNQYEVQAYRDQLTILGYQEVLDSEIPADICIINTCAVTA
SAESSGRHAVRQLCRONPTAHIVVTGCLGESDKEFFASLDRQCTLVSNYEKSRLIEKIFS
YDTTFPEFKHSFEGKSRAFIKVQDGCNSFCSVCIIPYLGRSVSRPAEKILAEIAGVVD
QGYREVVIAGINVGDVCDGERSLASLIEQVDRIPGIERIRISSIDPDDITEDLHRAITSS GITHEVALAGING DICEOGRADHADI DE QUINTIPATERI RISSIDPIDI I EDIGRATISS RHTCPSSHLVLQSGSNSILKRMNRKYSRGDFLDCVEKFRASDPRVAJTTDVIVGFPGESD ODFEDTLRI I EDVGFI KVHSFPFSARRRI KAYTFDNQI PNOVI YEKKYLAEVAKRVGQK EMMKRLGETTEVLVEKVTGQVATGHSPYFEKVSFPVVGTVAINTIV SVRLDRVEEEGLIG

CPn_0478 557640 556210
hflx-gtp binding Protein
wHcgpldtidtpcecosossgnsicgarfdlprkeqdpsoalavasyonktdsovveehld
elisladscgisvletrswilktpsastyinvokleeieefilkeppsigtlideettps
Ognlekricolvvldatelileifssraltaeaniovolaoaryllprlkrlughlsrok
sgcssgofykoegekoieldrrwyrerihklsaolkavykoraerrkyksrrgiptfali
gytnsgkstllnlltaadtyvedklfatldpktrkcvykogrhvlltdtvofirfirklphtl
vaarkstleaafhedvllhvvdashplalehvottydlfoedstektekpriitvlnkvdrlp
cosipmklrllsplpvlisaktoegionllslmteimoekslhvtlnppyteyckffelc
dagvvassryoedflvveaylpkelokkfrppfisyyfpedcgdddegrgpvlessfgd

CPn_0479 558434 557616 CPT_0479
558434 557616
phnP-Metal Dependent Hydrolase
ATGMVRDIOSESIGKLVFLGTGNPEGIPVPFCSGRVCONTGIHRLRSSVLIQYONKTLVI
DAGPDFRTOMLVAGVSELDYIVFLTHEHYDHIGGIDDLRAWYIVTGRSLPLVLCASTYRFL
NKAKEYLFATPNVESGLPAVLEFTILNEDGGDEFFGGIPYTVYTYGRSCHVTGFRFGNL
AYLIDLGYDAKIFGVLDNVETLILGAGPYETTIFFCHKGSHUTVEEAKAFANHAGIKN
LITHIGHGLEAERDQHPEVTFAYDGMEYLWTL

559375 58650 CT38; hypothetical protein CONSTRUCTOR OF THE PROPERTY OF CPn_0481 56973 59339

No robust homolog present in Genebank/EMBL as of 11/7/98
SCLRIEGILMATSVPVTSSTSVGEANSENERTERTSRMYYAALVLGALSCLIFIAMIVI
FPOVGLWAVVLGFALGCLLLJGAIVFAVSGLVLGKTLEPSREATPPEIVACKEWTTOODV
LCNEYWRSELISLFLRCDLHESLIVTØSKDRSLDIDOSLONILKLEPLSTTLSLLKKDCVH
INTILHLVRCWNLLGVDLSPEYTAHAEELLLFLIEGOYYSPOLIKLIRYGDLACATSFEM
MARPALSENTANIVAN TEREFALIVAN AND MARPALSENTANIVAN TEREFE
EFELHERIELS AND FETAL VAN TEREFALIVES KILLINAMILJAWYON TAR JEKE
EFELHERIELS AND FETAL VAN TEREFALIVES KILLINAMILJAWYON TAR JEKE
EFELHERIELS AND FETAL VAN TEREFALIVES KILLINAMILJAWYON TAR JEKE
EFELHERIELS KILLINAMILJAWYON TAR JEKE
FETALBERIELS KILLINAMILJAWYON TAR JEKE CPn 0481 ERPLINENCE, LAPRO ESTA TEGANGHUDNUF TAFUSOVEAMENUFURTGUVVRVDHROTHEOTESNTETLENESUFUCS LYEYPLSYLIDMAVLLDCVPCTETSLEDQADYTVCLQGLDSMLSQFASRLQSGQKVLMPR DVLSEQAAVMLVHGLAAQGVSFQGLKALMYLTAVPQRMMLGALPLFESFPVFNRMKEFLG

CPn_0482 961764 560961
artJ-Arginine Peripliasmic Binding Protein
MLAYRAGTFMIKQIGREFRAFIFIMPLSLTSCESKIDRNRIWIVOTNATYPPFEYVDAQG
EVVGFDIDLAKAISEKJCKQLEVREFAFDALILNLKKHRIDAILAGMSITPSROKEIALL
PYYGDEVQELMVVSKISLETPVLPLITQYSSVAVQTOTFOEHVLLSQOGICVRSFDSTLEX
IMEVRYCKSPVAVLEPSVGRVVLKDFPNLVATRLELPPECWVLGCGLGVAKDRPEEIQTI
QQAITDLKSEGVIOSLTKKWQLSEVAYE CPn 0482

CPn_0483 561830 564964

No robust homolog present in Genebank/EMBL as of 11/7/98

IILIKKRAIFIBMFPIPPHCPPNNKNNFYHLTTDTKOPLLLÄILATIGYVLLHIITLGL

LLIHYYKHHÄVVRREGLETPPTLEKPESPENTIEILAKOPEKOEDKKPDVPKRGTPPPED

TPPPPPKAPSPASPKVPKOPADKKPTPPPEAPPPVRVATPMPLRPSSGOYWGCLNRMVS

MVLRRAPLPLPAMQVDPILGOFNPHFVASYPNRIDNEPMYFQIKQFKKIAQNPDLPOQMS

RLAQUSLGOALYHONDYLVNVPGGONCFYRAYAVGWLSALVEESSRNDIVFEQEATRLL

DLPFASSPANANLCAEMAELLQLCSTYCSFIDLYDGVILSOKHTATLIAFLRKLSAYAI

RQOIAASSNEETARALFISDMQDDLLPSVLEFLANNFYSELFONLIDNSALPYMGSNR

RQOIAASSNEETARALFISDMQDDLLPSVLEFLANNFYSELFONLIDNSALPYMGSNR

LYLLHLPHAPLFLTDAELGMSPEDQOLKQYEREIREAFAKLSRRIADSGWTERFNAI

VKDHMPEAIRCQYSRFLATIENRRSGDLPWSPALSFFAFLCTDESVFHKLCATFYKSLE

DII JASAPPQRSIGEILGISNASLSYLMEDLDSSWOREVISSNIMTILTHESLTLESSM

POLETLHKRIANLLKNVISTSFETPPLSNQPDLLSNLVNKLLVAIHSKLELKEHFNTVCS

ARSLRITRDESSGLSQEODLLYTQAVOLLFFILQHPOVNNRPETKDAVKELKHLLLFYCS

ARSLRITRDESSGLSQEODLLYTQAVOLLFFILQHPOVNNRPETKDAVKELKHLLLFYLL

ZYKKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYKKVENEKKLJOKLRSILGSLVLKPPARYPSTPSNKDKSTPSKSSPSTALS

DOFSKSFLIFCPLNYPYBLLJCKKTPLAARLDAFOREASHRFTOVKDKLLLSKYGFPLAT

ATINGYSRARDQLICNLLKNTVTASDGFCRSGFROSLIGYLHSLSSNELGDILLDDVKEQA

EANDVAAMTTYPLOFFAVCLINSDRDTVSEENIENFVAMHGFLNTISPERDARIFLIRFP

NHYGCLLPRNPRTEDONSKPDSSNP

565824 CPn_0484 CPH_U884
aroG-Deoxyheptonate Aldolase
RSELKTQLKSLVLHEVLILITFTYPLPRTLKQHPDEVHTVPISPNLSFGEGSPILIAGPC
RSELKTQLKSLVLHEVLILITFTYPLPRTLKQHPDEVHTVPISPNLSFGEGSPILIAGPC
TLESYEHTVSSALTVKEAGAQVFRGSIRKPRTSPFSFGGWEKECVLWHKEAQSIHGLPTE
TEVLDVRDVDVEITAEHVDILRIGAKWHNTPLLQEVSKSHRPIILKRSPAATLEEWLCAAE
YILASSPSCPGVILCERGIRTFEHSTRYTLDLNTVALLKEISSLPVIVDPSHAAGKRSLV
LPLASAGLSVGADGLMIEVHAHPEKALCDAKQQITPEELHLFAKKHFCPSESRAHAIS

565993 566229 CPn 0485 CT382.1 hypothetical protein
OPIGRTPTRVFLWRFMIKQACKFYLLQCLLCALYWLLKYCRKLLKGTLHHSEETLYQALL SSLIDLLYQLKQLPAPTNE

CPn_0486 567799 566405
hypothetical proline permease
AQHRSLLKGNIFHLGCGVLYFMNTSLFLFFLIAIQGICLYVGRRGSKKVEDRESYFLAGR
SLKIFPLMMTFIATOIGGGVLLGAAEEAFCYGYGGILYPLGVALGLIFLGMGPGKRLAEG
SLTTVVSIFEVFYGSKKURKIFAIFLSAGSLFFILVAQVIALDRLFSSFPFGKYVTVAFAF
VLASYTSTGGFRGVRTDVIQAGFLLIAVLVCGVSVMLSVPKSLSVLDPFQSLPCAKLSN
WIFMPMLFMLVEQDMVQRCVAASSFKRLQMAAVGAGLVLLLFNFIPLFLGSLGAKAGLKA
GPLIDTIAYFCNPSLAAVMAAAIGVAILSTADSLMNAVSQLIAEEYPTLKAPYYRVLVL
GLAVAAPLVAIGFTNIVDVLILSYSLSVCCLSVPVGFYYLLAPKGRRVSGAAAWAGVLVGA
LGYGZMGVJVSIGMGEGIL AWVGSLAAFESVJET FITMHKWAVKTOT CPn 0486 567799 566405 LGYGWVQIVSLGMFGELLAWVGSLVAFSFVGFIEITWKNKVKTQT

569833 568112 CPn_0487 569833 568112
CT384 hypothetical protein
RTTGGISLTYSSFRMASFRCYSLIFFCFCGSLFGSESLRYQLLIQDFAKVSEEGIGLLES
KEYSLLQAKLVLRALAQNSSFDDWFRSFKCQISYPELAHDRDVLEFGIQVLEGGIENF
KEYSLLQAKLVLRALAQNSSFDDWFRSFKCQISYPELAHDRDVLEFGIQVLEGGIENF
SVTVRAVSVLAIGLARDFRLVPELLQSCNDDSAIVRSLALQVAVNYGSESLKKAIVELAR
NDDSIHVRITAYQWALLQIEELLPFLRERAENKLVDSVERREAWKACLELSSOFLETGV
AKDDIDQALFTCEVLRKOMLPETTEIFTELLSVEHPEVQESLLLSALAWSHQLQNHKEFL
SKVRHVMCTSFFAKVRFQAAALHLHGDPLGRDSLVEGLRSPQPLVCEAASAALCSLGIH
GVPLAKEHLESLSSKAAAANLSILLLVSREDIERAGDVIARYLSNPEMCWAIEYFLWDAQ
WNLRGDTPFLYSDMIKREIGRKLIFLLAVARYSQAKAVTATFLSGQQAGWSFFSGMFWE
EGDVKTSEDLVTDACFAAKLEGALASLCQKKDQASLQRVSQLYNDSRWQDKLAILESVAF
SENLDAVPFLLDCCHHEAPSLRSAAAGGLFSIFK

CPn_0488 570147 569767 hita-HIT Family Hydrolase RKLPTCFAVNVTRSROHMTYFKOIIDGLIDGEKVFENENFIAIKDRFPQAPVHLLIIPKK PIPRFQDIPGDEMILMAEAGKIVOELAAEFGIADGYRVVINNGAEGGQAVFHLHIHLLGG RPLGAIA

571037 CPI_0489 571037 570096
CT387 hypothetical protein
RIVFATENYFOLVTNEDMIRRIVFMOTERSTGTHDQSFHADEVTAGALLITEDLVDENKT
IRSRDEVVLOKLENYGDVGAVGTENKRFOHHGVGYDGWSSAGMIGHYLKERGYMOGEE
YHFLANYTLVHGVDEQOWGREFSKEGFGSFIGTTLT LKTYNFREEEETNSDADFSCALHFTLDF
LCRUPKKFGYDRVGRTVREAMETEDMICTYFDP PLAMQENFFTLXGEKHFAAFVGFFSGD
QWILKGTPTNLDBRNDVRVFFFEIMAGLIGKELGKVSGTFGAVFGHKGLFLSVWTNRESC ORALRETEODROLL

CPn 0490 CT187 hypotherical protein
IMANUTHHIDVELOCHTENTER TEALER TOAKTOANTRE CLAQUIVEAVENHEELMARKEAFUMLEMEGLET THE LIMBELTER EIFOTT CYEET LY VROCYCALMMETHFORT CIPLLEFORNLEHF LYLEL IMBRLVVFLE LLIGT I CYEET LYG

95

23

FLPLMSKSIJTRPHLKIRKFLPL/GMVTDRPPV/PEDHKI

POGLRITAADILEPTTOESGDI/EFYGSTSEPIERIPI

POGLRITAADILEPTTOESGDI/EFYGSTSEPIERIPI

EPYKEHSFFFYRDMLQE

PRISPSDERHAREIQKH

LEDDPCPPFLKAMETOHITSOGVLFSRYFPSASLKOMFLSNYSRYYLOHIYFQIPSPTSG

EFFSNRDRSFLLDL/FAGISVFWADLESKRLLOY LKRNNKDYGMFVPKHQAEQFAQSYFI

GIHGSCLIAGDYDEFLRELLTCMHTLSQOFTIPEFPPQTPLAILTGCGSGAMELANNVAT

ELSILSCGNLIGLDTTNAYVEAKMSYAIPDLLEROADFHVDLAVFVIGGMGTDFELLLEL

ISLKTGKKALVPVFLIGPVDYWKSKITALYNSNHAVGTIRGSEWVHNCLFCLSSAKAGIA

CPn_0491 574595 573336
CT399 hypothetical protein
ILSSLYTVFTMKTAFHSCYSWFCWLFSFLVLFVGGIAGGEPLCPDCKYETKSVLRSDQLP
ILSSLYTVFTMKTAFHSCYSWFCWLFSFLVLFVGGIAGGEPLCPDCKYETKSVLRSDQLP
DHLMNYENDCYLTGYVQSLLDMHFLDSRTQVVIEKNRAYLFSLPVDSSLSEAITNFVRDL
PFICAVEICERPYGECITRSSAERPLLPKEKTLGMPIFCGKEGVWLPQNTILFSPLIADP
RQVTNSAGIRFNEKVVGNRVGATIFGGDFILLRLFDVSRFHVDCDFGIQGGVFSVFDLDH
PESGMVNSDFFVAGLMSGAIDKWSFRFRLWHLSSHLGDEFILTHPNFPRFNLSDEGVDLF
ISFRYTPQIRLYGGCGYIVSRDLTPPERPFYCEWGAELRPFGLREGNLHAQPIFAMHFRC
WEEQKFGLDOSYILGMEWAKFQEIGRKIRAVLEYHQGFSKEGQFIREPCNYYGFRLTYGF

CPn_0492 574643 574804 No robust homolog present in Genebank/EMBL as of 11/7/98 LFSLIFPICEERNSQOTYKHLHVESACFLLESPLKIHWSSPYGFPPFYRRDLKL

CPn_0493 575142 574855 No robust homolog present in Genebank/EMBL as of 11/7/98 SKTEGSHSKTSKGFVGRFVQWIRTFTGRGSKKRSPSSFSPTHPYIRLRTYTRSPKQSGVE RKQEDAETISFIETPKGILKKPGNKDPKGKHVHWKDS

CPn_0494 575370 575146
No robust homolog present in Genebank/EMBL as of 11/7/98
VINIRVNPYGSYRGRNPSFEDGKKDVPLSGNSRLHRRGGIRRKHKSASVGVTSGSKTGKA
SLEKKVKGISEAHFK

CPn_0495 575507 576793
aspc-aspartate Aminotransferas
RRLKKNQKMAIQKAGAFLRCLPSESRPYLEHAMRRNPHFSLLKPQYLFSEISKKLAQFRK
ENPEISVIDLSIGDTTQPLCRSITQAIKEFCVSQEKQETYRGYGPETGLEKLRYKIASEV
YENRISPEEIFISGAKPDIFRLFSFFGSEKTLGLQDPVYPAYRDIAHITGIRDIIPLAC
RKETGFIPELPNQSLDILCLCYPMPTGTVLTFOQLQALVMYANQHGTVLIFDAAYSAF
VSDBSLEKSIFEIPEAKYCAIEINSFSKSLGFTGMRLAMNVIPKELTYDINEPMINDMKR
LFÄTTFM3ASLLMQEAGYYGLDLFPTPPAISLYLTNAQKLKKSLETAGFSVHGGDHAPYL
WWEIPEGISDEEAFDFFLHQYHIAVTPGHGFGSCGQGFVRFSALTQPQNIALACDRLCTA
SLKETMYLA

CPL_0496 576751 577812
CT991 hypothetical protein
PPHMRFTKRNDGSCMTILRKLSGYLFFFSLFCSFIYVATCGSOPDSVSSPKIAIFLSFPH
PPHMRFTKRNDGSCMTILRKLSGYLFFFSLFCSFIYVATCGSOPDSVSSPKIAIFLSFPH
PLHEDCSKSCIETLKDFENLPEIVVLNAEDSIVKARKIARSLHTDKNVVAIVTLGTIATK
VMSAIETOKPVIYAAVPDRESLTLFKNTMNIYGVNDTLDINGYCFAIQAVATNAQSIVYL
KPSEPFSDLOKEIVKKLHASGIEVIEISITSSTFKTRIRQAIDKRPSAIFIPLSPLSHK
EGTATLOEILKEKIPIITDDTSLISEGACIACSVDYKKSGKQIAKIVHHLLYNNHDVDSL
RKITAORLSPTTTFNEDIIKYLGIKLHKTERNQFLSFKSKKLEKSEKGKNVAVS

CPD_0497 578107 577820
CT388 hypothetical protein

CPÄÖ497 578107 577820 CT388 hypothetical protein IFORVVLDDSWILEVKVTPKAKENKIVGFDGQALKVRVTEPPEKGKANDAVISLLAKALS LPÄRDVTLIAGETSRKKKFLLPNRVQDIIFSLHIDV

CPA_0498 579062 578085

No robust homolog present in Genebank/EMBL as of 11/7/98

YCRIRRAPFMNRKKARWVVALFAMTALISVGCCPWSQAKSRCSIDKYIPVVNRLLEVCGL

PEÄRNVEDLIESSSAWVLTPEERFSGELVSICQVKDEHAFYNDLSLLIMTQAVPSYSATY

DCÄVVFGGPLPALRQRLDFLVREWQRGVEFKKIVFLCGERGRYOSIEEQCHFFDSTATY

PTEEMWESGNRVTPSSEEEIAKFVMVMLLPRAWRDSTSGVRVTFLLAKPEENRVVANRK

DTLLLFRSYQEAFPGRVLFVSSQPFIGLDACRVGQFFKGESYDLAGPGFAQGVLKYHWAP

RIGLETLAEWLKETNGCLNISEGCFG

CPn_0499 580404 579205

No robust homolog present in Genebank/EMBL as of 11/7/98
LSYYLLIFYFCNGSTHSSWOGSGTPNPEEVTSPESTEENKNVVSSDEAQATHAVALPIV
TOLSLPPGVGTSSEETASNPRVDEIVAEVSSSRAVADOISSLVERVGELLDDLKGAQSLF
TSFOSELKNCLPAWKSSTRRLETRGAGDNADIARLELFRSDYEAVLGHANOFHGKAHLIL
SKLTDVHHKLOGLSREDDLSLAFDNDRVLEHLGSIGLDVDAEGGWNSLSCERGIPRLVLTA
DSMLVQIKKVNLPTVEELRTLQGTTESSSDPRVEESLSCCERLLNELRRLWANFVGFISS
CYDNIVFVLMWIVRRINLLEGLGCLFFINDPASQEDQRSSSGERSTRRERLSRRSDLSEE
EMIVRAEGESIHPESEPHGOGRNOPSRGDKODSDSEEETEL

CPN_0500 580547 582362
proS-Proly1 tRNA Synthetase
QPHSMKTSQLFYKTSKNANKSAAVLSNELLEKAGYLFKVSKGVYTYTPLLMRVVSKMMNI
IREELNAIGGOELLEPLLHNAELWOHTGRWEAFTSEGLLYTLKDREGKSHCLAPTHEEVI
CSFVAQWLSSKROLPLHLYQLATKFRDEIRPREGLIRSRELLMEDSYTFSDSPEGNNEQY
EKLRSAYSKIFDRIGLAYYIVTADGKIGKGKSEEFQVLCSLGEDTICVSGSYGANIEAA
VSIPPOHAYDREFLPVEEVATPOITTIEALAMFFSIPLHKILKTLVVKLSYNEEKFIAI
GMRGDRQVNLVKVASKLNADDIALASDEEIERVLCTEKGFIQPLNCPIDFFAUFTTSPMT
NFVCAGNAKDKHYVNVNMDRDLLPPQYGDFLLAEEGDTCPENPGHPYRIYGCIEVAHFIR
LTFPYTDSFEVNFODEHGOTQOCWMCTYGIGVGRTLAACVEQLADDRGIVWFKALAPFSI
TIAFHCGDTVSQELAETIYHELQSGCYEFLLDDRDERLGFKLKDSDLIGIFYKLLIGKSY
QUXIFFEIEDRGGEKYTVSPEAFPTWCQNHLA

 CPn_0502

GUPE-HSP-70 COLDICTOR

GDVMTDTPPENEEQHESN/ONENE/EHLQQEIVTLKTELKEKNDKYLMALAESENSRKRL

GKERQELMOYALENTLIDFLNPIESMEWALGFATOMSDDVKNWALGFNMILNQFKQIFEE

KGILEYSSIGQKFNPFLHEAVQTEETSEVPEGTILEEFAKGYKIGERPIRVAKVKVAKAP

CPn_0505 588471 589106
*3-methyladenine DNA glycosylase
RKRLLRKKERKEPRNVLQEHFFLSEDVITLAQQLLGHKLITTHEGLITSGYIVETEAYR
GPDDKAGHAYNYRKTORNRAMYLKGGSAYLYRCYGMHHLLNVVTGPEDIPHAVLIRAILP
DOCKELMIORROWRDKPPHLLTNGPGKVCQALGISLENNRQRLNTPALYISKEKISGTLT
ATARIGIDYAQEYRDVPWRFLLSPEDSGKVLS

CPL_506 589055 589840
CT-31 hypothetical protein
CPM_ISPIERREGKSFILNNLKLYSKETNAHFLISCRRIMKKYFITGLVILLPLAITIAI
CPM_ISPIERREGKSFILNNLKLYSKETNAHFLISCRRIMKKYFITGLVILLPLAITIAI
LYPIERREJTOPTUGLASEFFEKFSFYTKHRALLKFVLQIILLEGLFFATVLLGFILTRIMI
KELLSIYBKILHRIPIIKTVYKAAQQVMTIIFGSKSGSFKQVVMVFPFNANVQCIGLVA
GBAPTVCCTGEKEDDPLVTVFIPTTFNPTSGFLTLFRKSDIVFLDMKIEDAFKYIISCGV
USTPMACDESPLPDELHQDQGS

On_0507 589898 590122 CN21.1 hypothetical protein STYPOFPLSGEIKKFNIELFMTRMSKQARRRAKSPKKRKPKYAIVHPAPAPRIVYKLHT NALSTSDSIFIPKIG

CPn_0508 590133 590300
CT421.2 hypothetical protein
SRIMSRHRSYGKSVKGVTKRNVLKRFERVEVLRKLGRWNDSTAKKVTGLPKTPILK

CPn_0509 590299 590808
(predicted Metalloenzyme)
NKFVFLYGNFIRVTQEKIKIHVSNEQTCIPIHLVSVEKLVLTLLEHLKVTTNEIFIYFLE
DKALAELHDKVFADPSLTDTITLPIDAFGDPAYPHVLGEAFISPQAALRFLENTSPNQED
IYEEISRYLVHSILHMLGYDDTSSEEKRKMRVKENQILCMLRKKHALLTA

CPn_0510 590804 591973
tlyC-CBS Domains (Hemolysin homolog)
QLNMLHILLAIFCILLFLAFGLTQPSCHGSSKFLKTLNQRFFKDKGREYPPFPSAPTILA
TLLCILYGALGTKLYTLLPPKTAHKDLLFWPLYSLSALIAYGFLPPWISTKVPRETTAHL
RFLASVFOLGLFPLQLLFYRRRPNQOVRSSTSFQSQLSEALSAFDNLIVREVMIFKVDIF
ALPEETTLOEALJVLVSEEGYSRVPVYKKNLDNITGILLVKDLLLLYTSSHDLSQPISSVA
KPPFYAPEIKKASSLLQEFROKHRHLAIIVNEYGFTEGIATMEDIIEEIIGEIADEHDVQ
ENTPYKKIGSSWIVDGRMNISDAEEYFNLKIDHENSYDTLGGHVFHKVGAVPQKGMRIHH
ENFDIEIITCTERNVGKLKITPRKRKFHIIS

CPn_0511 592141 592488
rsbV-5igma Regulatory Factor
MSDIQKEEHGSTTIFHLHCKLDGISSPEVQENISQSLAAGSKNIILDCAHLDYMSSAGIR
VLLQSYHQVQOHSGKIVUTTVPKTIEQTLYVTGFLSYFKIFNTVDEAIQTLNKDGD

CPn_0512 592538. 594412
CT425 hypothetical protein
SLPLTMRRSVCYVNPSIARAGGISTWKFLYSLATPLPAGTKCKFDLAGSGKPTDWEAPAT
SLPLTMRRSVCYVNPSIARAGGISTWKFLYSLATPLPAGTKCKFDLAGSGKPTDWEAPAT
DLSOTRIVIYAEMPEGE I EATA I PVKDNPVPOF EFTLDYBLQVGETLT I VMGASPNHPQ
VDDAGNGAQLFAQRRKPFYLY I DPTGEGNYDEPDVFSMD I FGONTLKK LEI FTPSYVVNNK
RFDITVRFEDEFONLTNPSPEETR I ELGZEHLRENLIMOUF I PETGFVI LPNLYFNERGI
RFI CLKNLSTOET I FISAPI KICFADSAFILMMGLLEGESERVDSEEN I ETCRRYFRDRAL
NFYAGGGFENGENLITDINKL I NOTVOLFNEEDF FITLOGFQYGGEPHLEGVRHILHTKE
TKSIGKHKEZKH I LAKLYKSTVNHDM IS I PISTTASKEHGFDFENFYBEFENVEL YNAW
SISHETTAALINNPPT LOKKOSEDHRITV I EGLK WILLRFGFVAGGLDDRGIYKDYFDSPQVQ
YSKOLTAA I I CHKYTRESLAFAHLEGARHEGYATTOEPF I VLSFHITGAPMGSELSTGSK PGLAV
NRIHESHVAGTALLKTVE I I RNGEVLHTFFFDGHHLDYEYDDMVPLGSVTLKDPHGKAPF
VFYYLRVTGADNAMANESTIWGIL

CPn_0513 594515 59575;
Fe if oxider eductions
Fe if oxider eductions
FEERLANGERISSEDALIILLLITHNEIGETVIGHLER IVHGCVHOMTTCLIOPPEKTSPLYG
FEERLANGERISSEDALIILLLITHNEIGETVIGHTRIFADQVPROPVGDTVYYGGTLYLYPTNE
CDERCKEGGFYAKTGDPKGMLYGPDDLIOQTGHTTPTTEVHTVGKGFPGKNLAYYYGDLF

TK IKEYDI'Q IH IKALTA IEYAYLSDLUNLS IRDVLLT
FLAPKRUSSDFLN IHKMANQLG IHSN ITMLLYHKEG
FILLKFAQENNVLGKRLRKSGCHA I PLKSLMAVAR I P. SMKALMYYLG I EAALDL
LSCGANDLSSTHIMGEKVFQMASSKEP I KMDAEGMAALI TQQGRTPCLTNSSHV

CPn_0515 596450 597181
ubiE-Ubiquinone Methyltransferase
EKNTTKALKNSONIMEPSTNKPDCKKIFDSIASKYDRTNTILSLGMHHFWNRSLIQILGS
GYSLLDLCAGTCKVAKRYIAAHPQASVTLVDFSSAMLDIAKQHLPQGSCSFIHSDINQLP
LENNSYPLAMAYGLRNLSDPHKALQEISRVLMPSGKLGILELTPPKKTHPTYSAHKLYL
RAVVPMIGKSVSKDPDAYSYLSKSIQQLPKDHDLEDLFSKSGFYIAKKKKLFLGAATIWL

CPn_0516 598904 597255
NO robust homolog present in Genebank/EMBL as of 11/7/98
NISISFRVSWFVKIILAVLGRAIAKAYYYCMYARGLCDFPTLVPNERLPIGFFVPOHTS
GAKGKEFAKRNFSIISGLDDILKLCILQRRPFALQWONLSVKSDYEEAGPAIGIRSLEPQ
VSQISPAHGRLCSTLVQMAPILGSEEDLWWLEETIKKIKFPKSLGSKDAVIVDSEMVPVN
ANPYQEIPAASETVESSPVAPGNTTDTMPAASGTTDTTSGVSEAAAAEAAVDSTFGTEEE
PSFSLRYALVVQNVPYPEPPKEPFVMFTDEEKSLILEATRARMELDLYNGYLADYELSK
DEIQKHVPDLPENMTRINRWSERLYKFFFKTKKEGLEEIFLNKELGNMILARGLAATQSQ
ARIKVFNSLVAMLLQSFNVGRSCTAKPLPTSKLDLFKSEFFSKPKNNILTEFLVASDEEI
LFKGLRVLEEGIEGWYDHPDQAGEIRSVLEGLVQAGRISGYWENDPGRFVLLRGVGERT
ELVELLESLVASGEIMQFFESSDEEGAFIIDNEPSKTAMLKQRFKSCVRTKLVGSFADES
LPRGRFTILV

CPn_0517 599637 598795
No robust homolog present in Genebank/EMBL as of 11/7/98
FIMSSLLSGGRIEPTRVTCSLKTYLEDTSQNQLSTRLVRASVIFLCALLIILVCVALSSL
IPSIMALATSFTVMSLLIFVMSLLGDVAIISYLTYSTVTSYRONKRAFEIHKPARSVYYE
GVRHWDLGRSSLGTGEIFIVRTLFSPFQNHGLNHALAAKIFLFMEHFSPEPPNEPLVDWA
CLIRDFRPHVSSLCFVIEKQGSSLRYKEGNTICEAFRSDVDAHFAMVDCYRLIHSKLIIE
KMELKVIDIIPSVMVREDYPSRPGEGYREGLLRWYGGKGAL

CTM_0518 600806 599832
CTM_0518 600806 599832
CTM_0518 600806 599832
CTM_0518 FMTYPVPONPLLIAILRILADAFSKSDDERDFYLDRVEGFILYIDLDKDQEDLNKIYQEL
FMTYPVPONPLLIAILRILADAFSKSDDERDFYLDRVEGFILYIDLDKDQEDLNKIYQEL
EZNAENYCLIPKLTYFVKKIMETFINEKIYDIDTKEKFLEILQSKNAREQFLJEYYOPL
AELEKWQOFYVERSRIRIIEMLRNNKFHFVFEEDLDFTKNVLEQLKIHLFDAWGKEJTQ
AEQLLSNKAKIYYSNEALDNPPKRGRPFKQSAKVETETTISSDIYTKVPQAARRELALES
ITEPSSITFSEKFDTEEEFLANLRGSTRVEDQLNLTNLSERFASLKELSAKLGYDFJSTG
DFFGDDDEKVVTKTKGSKRGRKKSS

CPRE-0519 601707 600904
dåpf-Diaminopimelate Epimerase
OPFKLRILVYWMAFYSPSTISKYFIVSGAGNRFLLGETLPEVEDVRFLCQETR/DEFLYLV
KPŠSCADAQLIIFNSDGSRPTMCGNGLRCAIAHLASOKGKSDISVSTDSGLYFGVFISM
RVLVDMTLADWRASVHRLESRPDPLEREVVCIHTGVPHAVVILPEISTLDLSTLLGPFLRY
HÖTFSPGDVNVHVOILHCGLRVRTYERGVEGETAACGTGALASALVVSNSYGWKESIQ
IBENGGELMTVSQNRGRVYLQGSVTRDL

CRE-0520 60223 601646
clpp-CLP Protease
egenthadgevhklrdiiekellearryffsepyteksasdaikklw/Lelkdpgkpivf
vinspegsydagfavwdoikwltspyttyvtglaasmgsvlslcaafgrefatphsrimi
h@BSIGGPTTGQATDLDTHAREILKTKARIIDVYVEATNQPRDIIEKAIDRDMWMTANEA
kDBGLLDGILFSFNDL

CPE-0521 603803 602241
gipa-Serine Hydroxymethyltransferase
KSLLKVFEKFKKFAIVEIFTKVVAVVSLHKKILENASCKKGOSLASTAYLAALDHLLNAF
PSIGERI IDELKSQRSHLKMIASENYSSLSVQLAMGNLLTDWYCEGSPFKRFYSCCENVD
AIEMECVETAKELFAADCACVQPHSGADANLLAVMAILTHKYGGPAVSKLGYKTVNELTE
EEYTILKAEMSSCVCLGPSLNSGGHLTHGNVRLKMMSKLMJCFPYDVNPDTEECFOYAEIS
RLAKEYKFWYLIAGYSSYSRRLNFAVLKQIAEDCGSVLWYDMAHFAGLVAGGVFVDEENP
IPYADIVTTTTHKTLGPFGGLVLATREVESTLNKACPLAWGGPLFHVIAAKTVALKEAL
SVDFKKYAHQVVNNARRLAEFFLSHGLRLLTGGTDNHMYVIDGSLGISGKIAEDILSSV
GIAVNRNSLPSDAIGKWDTSGIRLGTPALTTLGMGIDEMEEVADIIVKVLRNIRLSCHVE
GSSKKNKGELPEAIAQEARDRVRNLLLRFPLYPEIDLEALV

CPn_0522 603825 604655
CT433 hypothetical protein
REPLSPEKTSLAFKVKNVNQRMIKKNQGKKKNYFØYIPLKVQKLROPSFYPKRLMTLYLG
LNOKTARKYQAHYLPILTLFPYAKSTPONKRALØFLPQATHVILTSPSSTHLFLSRMTSL
LSYATLKTKTYLC IGESTKERLLSFLOOVKYVÅROEIABGIFFLLQALPSSARILYPHS
SLARPVIREFLYNRFTFFSYPHYTVKPRKLKKNILSKYKKIIFTSPSTVRAFAKIFPRFP
EYTYWCQGRMTLQEPQKFSSQKOVSLLETLGKSRTSP

Ctn_0523 604720 605052 No robust homolog present in genebank/EMBL as of 11/7/98 PMAGSATFGFDCTAFSLEFPATREYNEFMLALFVTIAIALWHALLATTIAIGLCIHPLC SFIFUTAIFLYFISRYICSHYARNVYIALDVVPDHOKLODMRSHSPIFGDR

CPn_0524 605079 606179

By robust homolosi present in Genebark/EMBL as of 11/7/98
PYFYERIVMCOPORTECTOVIVES/PROKETAPEVOFTIAK OTTATLAGGALGALVAG
FOUTIVI SHOVELALLOTTALEIS/VTELVYHOMTIE/YOSMOVYLEONER PLIKAMORKN
VIJ YOSHMOFYNNILLOTTALEIS/VTELVYHOMTIE/YOSMOVYLEONER PLIKAMORKN
VIJ YOSHMOFYNNILLOTTALEIS/VTELVYHOMTIE/YOSMOVYLEONER GENPOPPINL
LIJITATULET LLYTELKIKKSVMPTOKOPROGORAGEDEFOSTEVRVVKLENEALDOTFINL
LIJITATULET LLYTELKITVOSTEKEELPINOGEYYROALIAYENICKAALECHAALVALDLE
TODYZEVIPEELLIKESTEYYMDYTOAPOKVALLOA FOITALEYPORSLLVILQOPFINTLE
TOOYZEVIPEELLIKESTEYYMDYTOAPOKVALLOA FOITALEYPORSLLVILQOPFINTLE
TOOYZEVIPEELIKESTEYYMDYTOAPOKVALLOA FOITALEYPORSLLVILQOPFINTLE
TOOYZEVIPEELIKESTEYYMDYTOAPOKVALLOA FOITALEYPORSLLVILQOPFINTLE

CPn_0525
CT198 hypothet ical/f
GIIFMIDALLSILAJOEDIYMIRLMRVKKEHQKELAKVOSLKSDIRRKVOEKELEMENL
GIIFMIDALLSILAJOEDIYMIRLMRVKKEHQKELAKVOSLKSDIRRKVOEKELEMENL
KTOIRODENRIOEISEGINKLENOOAAVKKHDEFNALTOEMTTANKERRSLEHQLSDLHD
KQAGGEDLIVSLKESIASTENSSSVIEKEIFESIKKINEEGKALLEORTELKHATNPELL
KQAGGEDLIVSLKESIASTENSSSVIEKEIFESIKKINEGKALLEORTELKHATNPELL
SIVERLLANKKDRVYPIENRVCSGCHIVLTPOHENLVRKKDRLIFCEHCSRILYWQESQ
VNAQENSTAKRRRRAAV

THE MINDS 57 CONTROL AND THE STATES OF THE S

CPM_05¢7 609910 608726
sucB-pihydrolipoamide Succinyltransferase
sucB-pihydrolipoamide Succinyltransferase
swyHyserrepkigetssggsivamiknigdhyardeplievstokiatelpspkagrlvr
fcvybdbevasgdviglieleeiseaddestsceptscetkseagssssswfspavi.sl
aprogicloniqkiagrokggrvtrodleavisesogvsipeifogevnripmsplrrai
asstsssbevehasluvddvutdunisserorfluthovklitisfivollaotiro
fyllnoslogitivkksvivogvavniksegvvvev ihncopaglysiakaladlssrar
lokklobsevogosvivtnfohtgaligheirypevailgigtiokrvvvrdddslairk
hvyvtlifdhrvldgiygsefltslknrlesvtmg

CPR_0528 611165 609921
gltT-Glutamate Symport
pMKLURMKIFIGLFVQVTLGLVLEDKAIFFKPIGDIFLNLLSMVVYPLVFCSMVLGIASIS
LMKLUMKIFIGLFVQVTLGLVLEDKAIFFKPIGDIFLNLLSMVVYPLVFCSMVLGIASIS
LMKLURGIGIKSVGLYLGTTALAIVIGLCFAWIFSPGNGCDFAQAQSMDSAVTVIDSNKT
AAYFLSIIAQVFPSNPVRSFAEGNILQIIIFAIFLGIALRLSGERGPVERFIDGFSEIM
LRWNNIMSFAPYGVGASMAWISGNHGLGVLWQLGKFIIAYYLACLFHATLVFGGLVAFG
CKMSFSKFLSSMDAISCAVSTASSSATLFVTMRCVSKNLGVSAEVSGFVLPLGATVNMN
GTAIFCGMAAVFIAQAYNCPLSLSSLLLLUVTATFSAVGSAGVPGGGNITLGSVLASVGL
PIQGIAILAGIDRLRDIVGTPMNILGDAVVATYVASGEGELSPYESIKQESVETT

CPR_0529 612298 611165
ycaH-ATPase
FSCKEIRAFKRITMKKRFPSTLFLFYRRVTIAISLEGILGMGWLGSLLSKVFAFLVACWN
FSWSTPYRARSTVISVGNIVVOGAGKTFTVLWILAEALRILRGYSCOVLSRGYKSOSSROK
KLITVVDSKVHSASYVGDEPLLHAEKLPEDSSWVHKDRRISAARAAEKFGILLIDDGLOYR
KLHKDVEIAVVNGODPLOGRAFFPKGRLRDFPLELKTVDAIIVNGGKEAGTVVKRVSNA
PQIFVKPTIASVWTHNGERIPKEALRELRVCVFCGLGFPCGFLNHLREEGIHILGKYLL
PDHAAITKKELNYFCQMAMRQCQGLLCTEKDSVKLPRLSGEVSLLPIAKVEMRLSVNQD
DTLSLLNMIEQIHNRGN

CPn_0530 613323 612460
spoil-rna methylase
svvlmgkflmrrcsslafwefcsmdcigkhnplvkealalkrsrcrksswflvegareig
kalrtgylloqhyfcsthlsekekeflyfliknstrkilycldstlaglsfkehidsfyau
gkrumnkepfligrknagpfylliegvekrgnvgailriadgagvogvilchpiydlyn
kvyrsslgafyfspilsiskebkelfkgesmyfytspraftmyfsknylgptalvfgs
ekdgltedwfsedfseialpmlgesdslnlatsvaavayevvrgrwyn

CPn_0531 614198 613245
SAM dependent methyltransferase
DSSKDDFRKEKGRRKSQYRDRYVNKDTOGRHSKTYFSLIRERLVMDYKLLDSGDGNKLECF
DSSKDDFRKEKGRRKSQYRDRYVNKDTOGRHSKTYFSLIRERLVMDYKLLDSGDGNKLECF
GPVTLIRPSSIAVMPKSRPELWSQAQLQYVREGERGAWKNFKRLPEEWEVAFSDVRCLIK
RTPFGHLGVFPELMSFWPALKQAIEKHKERQVLNLFAYTGAGSIFAAKCGARVTHVDASQ
AAVRWAQRNVEKNAFPERFIFWVIEDVISFLKKEIRRNKKYQVLLDPPSYGRGPDGEVF
KIDKDLFFLLSLCSKLLADDASYFLLTSHTPGHTPEFLRAIARRSVPTLVSEAWSCGESF
CGEGVGALPSGSFVQWIA

CPn_0532 614716 614075
ribC/risA-Riboflavin Synthase
eSFCCKDSYVKWCGMFSGIIQELGEVCFFEAQGNGLSLGIKSTPLFVTPLVTGDSVAVDG
VCLTLTSCNESKIFFDVIPETLACTTLGEKRCSDQVNLEAALKWGDSIGGHLLSGHVFGT
AEIFLIKENRYYFROSKELSQYLFEKGFIAIDGISLTLVSVDSDTFSVGLIPETLQRTTL
GKKREGERVNIEIDMSTKIQVDTVKRILASSGKD

CPn_0533 614918 615385
CT406 hypothetical protein
EVAPMCCPFCNNGELKVIDSRNAPEANAIKRRRECLKCSQRFTTFETVELTLQVLKRDGR
YENFQESKLIHGLNAASSHTRIGQDQVHAIASNVKSELLGKQNREISTKEIGELVMKYLK
KADMIAYIRFACVYRRFKDVGELMEVLLSATPDMEK

CPn_0534 615389 615784
dksa-Dnak Suppressor
LNFTRSKVPLSDDEIEDFKKRLLEMKAKLSHTLEGNAQEVKKPNEATGYSQHQADQGTD
TFDRTISLEVTTKEYELLRQINRALEKINESSYGICDVSGEEIPLARLIAIPYATMTVKA
QEQFEKCLLSGN

CPn_0535 615763 616296
LspA-Lipoprotein Signal Peptidase
KRTPIMKLSSMATRERSTLLVITLEVLIDMYTKLVVLLQYKDLQILTHPTLYTHSWGRFS
FSIAPVFRGAAFGLFSNYKYFLFLLRIFVILGLLAYLFFKKKSIOSTTQTALVLLCAGA
IGNVGDIIFYGHIVDFISFNYKGWAFPTFNVADVLISLGTLLLVYKFYFPTKQTEKKR

CPn_0536 616300 617691
ddgA-f-Ally/Oly Permeuse
PGCLGER/RYFEKTMRRELISLESVEDDFFWSY/AFILITYLGVSFSWKSRFFOFTKFSQ
PGCLGER/RYFEKTMRRELISLESVEDDFFWSY/AFILITYLGVSFSWKSRFFOFTKFSQ
PGCLFF/YSQNFOGERETKQ-VHELKVFFASA/SHITGIGIF/AGIVTAACIGGTGALFWWII
AGITGSTVKYSEVYLGIKFRKLDRDGVY/OSGPW/FLIKAFKTPW/VIVAILLIYTYGVET
VGESV-YTGALREWNI-KKVYFMCGLLFLVFYATRGGLQR IOKIGSTVLPFFHILIYCALSL
YTTAKFFFTLFHILISTVFSGAFYGOSALOGFAOCTVATTTHQGISRAAYSGUGIGFGASI
TUSKFFFTLFHILISTVFSGAFYGOSALOGFAOCTVATTTHQGISRAAYSGUGIGFGASYF
HMVKFFLIAFFFVANYTTI ISYFLVGKKCAKFLYGNTGAKTYTLYGILITLPLFCFLSQNF
ALLIMOVSGALLIGFNLLGVFTLIKKEVTFPARAASLTETSLSTE

CPn 0537 617815 618189

CTRIA, I hypothetical protess LIFELEMBNYLLOGLIFCCVLLDIGMCTIFVMTICFLE LAPLMLAKKLLCCWLKKKKNRGGLSEDIDELLDEKKOR:

ILDOĞİKWCAALVLIWKV

618511 518128

CTB14 hypothetical protein TREINFACHWOSEGKCFVIKIKTERDIYMFRNNHKPKKTKCKRFRWLRGVLFGGFIATLL TREINFACHWOSEGKCFVIKIKTERDIYMFRNNHKPKKTKCKRFRWLRGVLFGGFIATLL

CPn_0539 618678 621545

pm_19-polymorphic membrane protein
GYNLLGLRHMKQMRLWGFLFLSSFCQVS/LRANDVLLPLSGIHSGEDLELFTLRSSSPTK
TTYSLRKDFIVCDFAGNSIHKPGAAFLNLKGDLFFINSTPLAALTFKNIHLGARGAGLFS
ESNVTFKGLHSLVLENNESWGSVLTISGDLSFINNTSVLCONNISYOPGGALLLLGRKSK
ESNVTFKGLHSLVLENNESWGSVLTISGDLSFINNTSVLCONNISYOPGGALLLLGRKSK
ESNVTFKGLHSLVLENNESWGSVLTISGDLSFINNTSVLCONNISYOPGGALLLLGRKSK
ALFFRDNRGTILFLKNKAVNQDESHPGYGGAVYSSISPGSFITFADNQELLFQENGELGG
AIYNDQGAITFENNFOTTSFFSNKASFGGAVYSRYCNLYSQMGDTLFTKNAAAKVGGAIH
ADYVHIRDCKGSIVFEENSATAGGAIAVNAVCDINAQGPVFINNSALGLMGGAIYMQAT
GSILRLHANQGDIEFCGNKVRSQFHSHINSTSNFTNNAITIQGAPREFSLSANEGHRICF
YDPIISATENNNSLYINNGRLLEAGGAVIFSGARLSPEHKKENKNTSI INDPVRLCSSDPAE
IRATEKASIEISGVPRVYGHTESFYENHEYASKPYTTSI ILSAKLVTAPSRPEKDIQNL
IIAESEYMGYQYQGSWEFSWSPNDTNFKKKTIILSAVPTYGEFSLDPKRRGSFIPTTLWSTF
SGLNIASNIVNNNYLNNSEVIPLQHLCVFGGPVQIMEONPKQSSNNLLVOHAGHNYGAR
IPFSFNTILSAALTQLFSSSSQNVAADKSHAQILLIGTVSLNKSWQALSLKSSFSYTEDSQ
WKHVFPFYKGTSRGSWRNYGWSGSVGMSYAYFKGIRYLKMTPFVDLQYTKLVQNPFVETG
YDPRYFSSSEHTNLSLPIGIALEMRFIGSRSSLFLQVSTSYIKDLRRWPQSSSASLVLNH
YTWDIQGVPLGKEALNITLNSTIKYKIVTAYMGISSTOREGSNLSANAHAGLSLSF 618678 621545 CPn_0539

CPn_0540

621631 626862

pmp_20-polymorphic membrane protein
fihliylsliefynisdrfssmkwlpatavfaavlpaltafgdpasveistshtggdpt
sdaaltgftosstetdgtytivgdtffstfthipvpvvtpdandsssnsskggssssga
savvaadpkggaafysneangtlftftdsgngsltunkmtgddaatyskgplyftgl
knltftgnesoksgaafysneangtlfttdsgngsltunkmtgddaatyskgplyftgl
knltftgnesoksggaaytegalttgistgfisinkasgggatyvkeatlfnaldsk
kfekntsggagggiytestltisnitksiefisnkasvpapapeptspapsslinsttid
tstlotraasatpavapvaavtpfiststdeagatyakggisistfkdltfksnsas
vdatlfvdsstiegsgaifaabsigioqottgtstrstnaksggitavogvtledla
nildtnntckgegaitytkkaltinngallttfsgntanksggitavogvtledla
nildtnntckgegaitytkkaltinngallttfsgntanksggitavogvtledla
nildtnntckgegaitytkkaltinngallttfsgntanksggitavogvtledla
nildtnntckgegaitytksaltinngallttfsgntanksggitavogvtlutlit
schilgenssekhggglslasgksitmtslesfclanktakenoganvpenivltfty
tokelfoenssekhggglslasgksitmtslesfclanktakenoganvpenivltfty
tokabktocsftyitnvnitnntatgnggiagkahfbridnivpsnoakkgggvtl
edaellekvitgsvsontatesogitakdiolalegstitdnvognakkgggvtl
edaellekvitgsvsontatesogitavdiolalegstitdnvognakkgggvtl
edaellekvitgsvsontatesogitavdiolalegstitdnvognakkgggvtl
edaellekvitgsvsontatesogitavdiolalegstitdnvognakkgggvtl
edaellekvitgsvsontatesogitavdiolalegstitdnvognakkgggvtl
edaellekvitgsvsontatesogitavdiolalegstitdnvognesultstnings
saatkktstkolaggaifsaavtiennsphifinnsakseattaatargkdscgali
nigetidisrtstolagaifsaavtiennsphifinnsakseattaatargkdscgali
nigetidisrtstolagaifsaavtiennsphifinnsakseattaatargkdscgali
nigetidisrtstovtilssenssiffknnctatnkycsiannsfaatistessakkai
nigetidisrtstovtilssenssiffknnctatnkycsiannsfaatistessakkai
vsprindskoklditottvilldpronlynnsykgedritlfninnavappikl
vsprinadskoklditottvilldpronlynnsykgedritlfninnsasgavtannvtlo
gensteldvbrefssacsygnlaipersvogalawreiltynksvsavlpyllennpra
typelstkekgnvynvlptrnaaraevssojylgsywtlygtytidasmtlvonangei
fläaafsovfghaeseyhldnykkkgsghstoaslyagnifyfpalrsphillennpra
teplstkekgnvynvlptrnaaraevssojylgsywtlygtytidasmtlvonangei 621631 626862

627137 628003 CPH_D541
627137 628003
SOLUTE binding protein (--yebL-Synechocystis Adhesin Homolog)
NNRSSYCTAFVMHKVIVFIFLTLYSLKSYGNDVIDKPHVLVSIAPYKFLVEQIABETCFV
YA LINNYDHTVELPPOQIKELRQGDLWFRIGEAFEKTCERNLTCQCVDLSQWYSLLQG
KYCTNQHTTNYDTHTWLSPKNLKVQVETIVTTLSKKYPQHATLYSSNGEKLLLJLDQDLW
EILTITSKAKQRHILVSHGAFGYFCRDYNFSOHTIEKSSHVEPSPKDVARVFRDIEQYKI
SSVILLEYSGRRSSAMLADRFHMHTVNLDPYAENVLVNLKTIATTFSSL

628737 628000 CPN_0542 628000 62873/
ABC Transporter ATPASE
FMTIRILAEGLAFRYGSKGPNIIHDVSFSVYDGDFIGIIGPNGGGKSTLTMLILGLLTPT
FGSLKTFFSHSAGKQTHSMIGWYPQHFSYDPCFPISVRDVVLSGRLSQUSWHGKYKKKDF
EAVDHALDLVGLSDHHHHCFAHLSGGQIGRVLLARALASYPEILILDEPTINIDPDNOOR
ILSILKKLNRTCTILMVTHDLHHTTNYFNKVFYMNKTLTSLADTSTLTDQFCCHPYKNQE

628710 629603 CPn 0543 CPR_0543
(Metal Transport Protein)
KSGIFMLSSLIRDSFPLLILLPTFLAALGASVAGGVMGTYIVVKRIVSISGSISHAILGG
IGLTLMIQVKLHLSFFPMYGAIVGAIFLALCIGKIHLKYQEREDSLIAMIMSVMMAIGII
FISRLPTFNGELINFLFGNILMVTPSDLYSLLIFDLLVLGIVLCHTRFLALCFDERYTA
LNHCSVOLMYFLLLVLTAITIVMLIYVMGTILMLSMLVLPVAIACRFSYKMTRIMFISVL
LNHLCSFSGICIAYCLDFPVGPTISLLMGLGYTASLCVKKRYNPSTPSPVSPEINTNV

630598 CPn_0544 CPIL_0544 6305-98 629525
yhb2-otp-binding-protein
REMPFORTINGEFECKKOKNV IMPUDEITLELRAGK/ANGVVAWRKEKYLPKOGPYGGN
REMV ROWI LEATTOVY:FEAYRM LEPLKAPEGGGATTMRTGROCKDL LVSVPTOTLLRD
AETGELLHDETVOGERLLUNG/SCK/XIKGNTFFKTSVIRAPTKATFGKRGETRGVELELKL
FADIGLIV:FENA-KROTLFRITAHTEVKVOAYPETTILAFSLGUVLCKORLYGKPMIJADI
HILIFANIGNKOD/SDFIRH/ERTLLLLFV/IDV/SKRERNSPEEDLETLIHELHSHQPDFEK
ROMINALINK/IDD/SIPEGER/SCEPCYRFF/L/SGGGVUX/SYRFFTORLAV

CTD_0545 630008 630633 11.27 LEP ribodomat protein TEAMERIPAMARKEQUADRICROPSKRECZKOGNOGKVETCHILLARGROTEMIRAQNVG RIGHARLALVICTVVMKKTNETYTUVVEGO

430112 CPn 0546

26

43158 632188 _0547 yobB family ALMPRONEZPLEFORWINGTHI U.C.HRELEGSTRET FHALMALSUTING H. WVAPECU, TROUTDEGING / PHATE BATES THE FILE WADES TO STATE STATES AND ASSESSED TO SECURE SERVING THE SECURE EALKSLKPNGKISHVAITIE FGCGDGVQCFC'/LTVMEYC

33234 632191 CPH_0548 531234 632191

cysj_Sulfite Reductase
cysj_Sulfite Reductase
cysj_Sulfite Reductase
cysj_Sulfite Reductase
cysj_Sulfite Reductase
cysp_Sulfite
CPn_0549 633662 633255
rs10-s10 Ribosomal Protein
PODVOHOPHNOHSLLRFLKKFKKRLLRSKGCMKQQKQKIRIRLKGFDQGQLDRSTADIVE
TARRTGARRVYDF1PLPTKREVYTVLRSPHVDKKSREQFEIRTHKRLVDILDPTGKTIDAL
KMLALPAGVBIKIKAA

CPn 0550/

635688 CPn_0550 635688 633580
fusA-Elphgation Factor G
fusA-Elphgation Factor G
fusA-Elphgation Factor G
fusYespikfmsnoepdlsairnigimahidagkttterilfyagrthkigevheggatm
dwmage/ergititsaattvfhigakiniidtpghvdftieverslrvlbgavavfdavs
gverosetvmroadkygyfriafynkkmdrmgadyfaavesmkeklganafpvhcfiges
gverosetvmroadkygyfriafynkkmdrmgadyfaavesmkeklganafpvhcfiges
gverosetvmroadkygyfildtigakweekeissdlkercaelranlleelatidesneafm
mkvi_edpdsitedeihqwmrgvienkinpvlcgtafknkgvgollnvivkwlpspldrg
nirginlktdoeisleprrdoplaalafkintdpvlgriffirjysgtlkkgsailnstk
kkeriskllemhanertdrdeffvgdigacvglkfsvtgdtlcddnoeivlerieffdp
viomaiepkskgdreklagalsslseedptravstneetrgotilsdmgelhldiladrmi
refkveanvgkpqvsyketitvsgnsetkyvkgsggrgcyahvlceiephepckgbev
fivogviekeyipavikgiedslntgvlagyglydvkvsivfgsyhevdssemafkicgs
mavkdacrkakpvilepimkvavitpedhlgdvigdlnrrrgkilggesrgmaqvnaev
plsemfgyttslrsltsgratstmepaffakvpgkigeeivkk

636174 635698 rs7-s7 Ribosomal Protein
mymsrrnsakknoipodpiygsvilekfinkvmmhgkksvarkivysalerfgkkinlen
vlegfgealenakpilevrsrrvggatygvpvevaserrnclamgwiikharskpgksme
vglatelidcfnkggatikkredthrmaeankafahykw

636698 636219 UPT_UDDZ BJ8698 BJ8219
ES12-S12 Ribosomal Protein
IQAGYVPSSSENKPLETKRALLYISMLVVVRLKREEYMPTINQLIRKRRKSSLARKKSPA
LQKCPOKRGVCLQVKTKTPKKPNSALRKVAWVRLSNGQEVIAYIGGEGHNLQEHSIVLIQ
GGRVKDLPGVRYHIVRGTLDCAAVKNRKQSRSRYGAKRPK

CPn_0553 637753 636812
No robust homolog present in Genebank/EMBL as of 11/7/98
GCMMRVVLRFLIFFILGRAVFPLRASESFSWETSTCLTVLGIPFIDIILTTNEDFVAQCG
LQIGTISSTNNAKIKEIFLIYKEKFPEASISFKRKEPLNLSQSHLSDLGILCMRNGETYA
EGMANKENGPALKQPKDLRLVLRCPNQPDTLLYSEKEAEKGIETHTTCLCNQSYTLLDOQL
LUYGDSIEKFLKETKRKNNHTLVDLCDSQVVTTFLGRFWSLLNYVQVLFLSEDSAKILAG
IPDLAQATQLLSHTVPLLFIYTNDSIHIIEQGKESSFTYNQDLTEPILGFLFGYINRGSM
EYCFNCAOSSIGET 637753 636812 EYCFNCAQSSLGET

CPn_0554 637806 638141 CT440 hypothetical protein VFSYLLLCIILVYVRFMYEGKSFMASPTPGQLHLQQKVESKAYDYSRSLAMIATALLFFI VALILSGLSLLPQVFLPFSGAYFIIGSFLAFIALGILLINCVCDLKQYLTSS

CPn_0555 638298 640241

tsp-Tail-Specific Protease

MFVMKKLVRLCVVLLSLLPNVLFSSDLLREEGIKKMMDKLIEYHVDAQEVSTDILSRSLS

SYIQSFDPHKSYLSNQEVAVELQSPETKKRLLKNYKAGNFAIYRNINQLIHESILRARQN

RNEWYKNRKELVLEASSYQISKQPMQWSKSLDEVKQRQRALLLSYLSLHLAGASSSRYGW

KEEQLAALCLRQIENHENYYLGINDHGVAMDRDEEAYQFHIRVVKALAHSLDAHTAYFSK

DEALAMRIQLEKGMCGIGVVLKEDIDGVVVREIIFOGPAAKSGDLQLGDIIYRVDGKDIE

HLSFRGVLDCLRGGGHGSTVVLDIHRGESDHTIALREKILLEDRRVDVSVEPYGDGVIE

VLYVYSRYADGTMKCYRTVSPKKFYDGPLAILVSKSSASSAEIVAQTLQDVGVALVVGDEQ

TYGKGTIGHQTITGDASQDDCFKVTVGKYYSFSGKSTDLQGVKSDILIFSLYAEDKLEGE

FLEHPLPADCCDNVLHDPLTDLDTQTRFWFQKYYLPNLQKQETLWREMLPQLTKNSEQRL

SENSNFQAFLSQIKSSEKTDLSYGSNDLQLEESINILKDMILLQQCRK 638298 640241

640921 CPR_055% Cysteine-Rich Protein
enomssnlhevgotorgaaapesvlmiveetaasgsvtaglgaitssrgmvnlligwakt
kfigpiresklfosracgitllvugillvvaglacmfifhsglganafwliipaaiglik
LUTTSLCFDEA-TSEKLMYFQKWAGVLEDGLDDGILMNSNKIFGHVKTEGNTSRATTPVL
NDGRCTPVLSPLVSKIARV

CPI_0557 642870 441144

omcB=60kDi_dybeging=frich_omp
E1fmerltrrvtvlaldemaccpacybieaavaeglitkivagaetkpapvemtakkvr
Lyrrrkcypveckorgabepokeryteepsyrcopveaqqeegcygrlysvkvnddxinvetcog
ypeyatvcepyt letlatokkocydvyttoqlppeaefyugopettipteidoklywktorloadorck ttwwkkelkeagcetaatvcacpelretrecopalctfreedoklyrrkcyberigaterytvontptypudiacogopyliptelodmepsyrchypvigalarnytvontptypudiacogopyliptelodmepsyrchypvigalarnytvontptypudiacogopyliptelodmepsyrchypvigaterytvontptypudiacogopyliptelodmepsyrchypvigaterytvontptypudiacopyliptelodmepsyrchypvigaterytvontptypudibyliptelodmepsyrchypvigaterytvontptypudibyliptelodmepsyrchypvigaterytvontptypudicytrypvigaterytvottelotropidacopyliptelodmepsyrchypvigaterytvottelotropidacopyliptelodmepsyrchypvigaterytvottelodmepsyrch

LGGDTUTGEVODTENTHVY

54303L 64 1 106 CPR_USDA omcA-9kDa-0ystoine-Rich Lipoprotein KLMKKAVLTAAMFCGVVSLSSCCRIVDCCFEDPCAPSSCNPCEVIRKKERSCGGNACGSY VPSCSNPCGSTECNSOSPQVKGCTSPDGRCKQ

E PRILIDONI 64 E POR 643927

TELLE DYBOT DE LI VAL E POLICIÓ
GRUPTILKURUS DEAPERTE DE LEUTOLE AND TELEPROY LLAÇUPTELANKARIOLE
FFLARGYENTOYNINL

645666 644098

CPN_0560 645666 644098

Q1tX-Clutamy1-trna Synthetase
RNSRFQCMKSLMSKOKRIMWENVRVRVAPSPTGDPHVGTAYMALFNEIFAKRFKGKMIL
RIEDTDRTRSRGDYEBNIFSALRKGIQMDEOPDVGGPYGPYRQSERTKIYYGYVETLLK
TDCAYKCFATPOELAEMRAVASTLGYRGGYDRYRYYLSPEEVASREAAGQPYTIRLKVPL
SGECVFEDYSKGRVVFPWADVDDQDVLVKSDGFPTYHFANVIDDHLMGITHVLRGEEMLSS
GECVFEDYSKGRVVFPWADVDDQDVLVKSDGFPTYHFANVIDHLMGITHVLRGEEMLSS
TPKHLLLYEAFGWEPPVFLHMPLLLMPDGTKLSKRKNFTSIFYRDSGYVRAFVNFITL
MGYSMEGDEEVYSLERIIETFNRRIGKSGAVFDIQKLDMMNKYLNHEGSPECLLKELQ
GKLINDEFFLKIIPLCGSRITTLAEFINLTSFFFSGLLEYRVEELLPQALSPEKAAILLY
SYVKYLEKTDQMTKETCYLGSKWLAQAFWYHHKKAIIPLLVAITGKKQGLPLFDSIEIL
GKPRARARLVYAEKLLGGVPKKLAATVDKFMQREDFEEATFDL

646407 645871

CPN_U551 04640/ 0435/1
euo-CHLPS Euo Protein
LMACEQHEGCYELEEREEIEDIKDSDTKWVSITQAAKLHNVTRQAIYVAIKQKKLKASKE
LMACEQHEGCYELEEREEIEDIKDSDTKWVSITQAAKLHNVTRQAIYVAIKQKKLKASKE
TRWEIDIKDLEEYKRNRYSRKKSLYQGELVFDNGKGCYSINQVAQIIGIFVQKVYATRT
GTIRGERKGAAWVIHVSEIERYKNEYLSKQAAKKLKGAEPKEHQAPNFEPPTEIFPESN

CPn_0562 648051 646918

*CHLPS 43 kDa protein homolog_1

NYKVIMSIAIAREQYAAILDMHEKPSIAMFSSEQARTSWEKRQAHPYLYRLLEIIWGVVK

FLIGLIFFIPIGLFWLQKICQNFILLGAGGWIFFPICRDSNLLRQAYAARLFSASFQDH
VSSVRRVCLQYDEVFIDGLELRLPNAKPDRWMLISNCHSDCLEFRTVLQGEKDWIFFRIAE
ESQSNLLIFNYPGVWKSQNITRNNVKSYQACVRYLRDEPAGPQARQIVAYGYSLGASV
QAEALSKEIADGSDSVRWFVVKDRGARSTGAVAKGFIGSLGVWLANLTHWNINSEKRSKD
LHCPELFIYGKDSQCNLIGDGLFKKETCFAAPFLDPKNLEECSGKKIPVAQTGLRHDHIL
SDDVIKEVAGHIQRHFDN

CPH_0563 650113 648293

rec_i_ssdna exonuclease
OYKMLLMDFSPKGPCGIKFMINSDNASAAGLLWAHPKEDPAFLGMIIKEFHLPPTVAQIF
ISRGFQTIQEIHKFLYSHLSSLYDPGLFLDMSKAVERLLLARDRKEHVMIYGDSDVDGMT
ISRGFQTIQEIHKFLYSHLSSLYDPGLFLDMSKAVERLLLARDRKEHVMIYGDSDVDGMT
GVÄLLVEFIRDIDVMVSYFFLGAILKQHGETSTLIAKLKEEGITLLITVDGSITAGKEVS
DIFRGGIDVIITDHHMPTGKIPHCVATLNPKLRDHTYPNRELIGVVAPKLARGVINALI
SRNIVPKSQCSLKKLLDLVTLGTITDVGVLLGENRVMVRYGIKELARGARPGLNKLCALC
GVEKSEVTSTDIVLKIAPKINSLGRLDDPAKGVELLLTQDDERVDALIMELDNINRERQR
IEÄSVFQDVQEILNSNPEILKQAAIVLSSTAMHARVIPIISARLAKTYNKPVVIIAIGRG
IGKSSARTIGSFPLLGVLKKCSSLLLSYGGHDFAAGVINKEDKVEDFKKKFYHLVNSSLK
KGPTLPHLEIDAYADFDAIDVDLLASMELFEPFGKONLMPIFYSKVRQVRYPKVLYGMHL
KLYLSQKERNLEGVAFGLGRHADALKASMHYPLEIAYTPRLSOTSGSGVIHLLVRDFRIS
SEPPRFSD

SEPRFSD

E
CPB_D564
654159
650145
SecD/SecF (fusion)
secDEsecF-Protein Export Proteins SecD/SecF (fusion)
secDEsecF-Protein Export Proteins SecD/SecF (fusion)
secDEsecF-Protein Export Proteins SecD/SecF (fusion)
secDEsecF-Protein Export Proteins SecD/SecF (fusion)
secDEsecF-Protein Export Proteins SecD/SecF (fusion)
secDEsecF-Protein Export Proteins SecDEsecPension
RESECTION
GTLSSLYIAPPLLLFMVRKENRSK

CPL_0565 655741 654533
CT449 hypothetical protein
NKLFCFLIFCFVNISAILFDSSFLLKIKRNSKRMLPSMKFPRISISDLIPTOMVIWMRGG
NNHYVPNANOLPKKILGOVLACFGLALLGCAAFAGVCQTIFPCIGLMILGLVLLGFAY
LOYSKGWSRFERPLFRETKVFEKPIMMLGCLSLIGSWKKIRPGCYYHPGCFQVEICEGSQ
EIVTKIFOKKGDRATSIFLIQEMDQIALRQGIFKSSLSRKTFAIDFSVVSSLLSIGREE
GYVLDFKVISMCSEQOASDRTHFKSAIYVNISGAAGEPGGRCYIDAYTKAFFTVLDQIGD
INIVKKHTIYVLTPILGVPDALPKEEQENLK/LISQAAFLYSAEQVAKRMREEKQDSIRIK
FIFTDPTSPTCLYFSPHHGSTPHSVTPISLSGFVGEQESYTFA 654533

650000 696890 CEn_0566

657817 odeA Phospharidate Cyridylyrransterase VLNSNKFKSKTGAYGDLFG

LVLTELVLLLYSSLFFLTSFALGFITATCGAVGTY
EYSSMAKAKMYPLSTFSALG

FLALSFLJTRWGHSLMGFFDALDATLLIVWVMSIF
RVRKSTIGALOLGSVTLFSILVYGIPIRLFLAVLYSFHMOEDFULJIWMASFLIATTKGA
DIFGYFFGKAPGNKKIAPQISPNKTVVGFVACCLGATLISFIFFLQIPTRFASYFPMPAI
LIPLGLALGITGFFGDIIESIFKRDAHLKMSNKLKAVCGMLDTLDSLLLSTPIAYLFLLI
TOSKFFFIG TOSKEFIG

CHA PROGRAM FROM YEAR OF THE THE TOTAL TOTAL AVAILABLE WATER THE RETAINMENT OF THE TOTAL AVAILABLE WATER THE TOTAL AVAILABLE OF THE TOTAL AVAILABLE WATER TOTAL AVAILABLE WATER TOTAL AVAILABLE OF THE TOTAL AVAILABLE WATER 657805 £58464

658/98 659099 CPn_0569 658/98 659099
plsc-Glycerol-3-P Acyltransferase
LFGFDNKTSSGENFSFTISKRANIFRICKFFTWVAFSLFYKLKVYGVKKNFIKGPAIIAV
HNNSFLDPIALHMCVHSCYTHLARASLFNIPWLWKQWGCFPVRQDEGNSAAFKIASRLFN
KRKKLVIYPEGARSPDCGLOPGKVGIGMMAAKSRVPIIPVYIRGTFEAFNRHQKIPHVWK
TITCVFGTPMYFDDIIGMPEIKNKETYQIITNQTMNKIAELKAWYESGCKGDVP

CPn_0570 659044 660789

args-arginyl tina transferase

TKLPSSKHGNINGGKETSFKLMSTLLSILSVICSQAIAKAFPNLEDMAPEITPSTKEHFG

TKLPSSKHGNINGGKETSFKLMSTLLSILSVICSQAIAKAFPNLEDMAPEITPSTKEHFG

TKLPSSKHGNINGGKETSFKLMSTLLSILSVICSQAIAKAFPNLEDMAPEITPSTKEHFG

HYCONDAMKLARVLKKAPRAIAEAIVAELPOEPFSLIEIAGAGFINFTFSPVGHDVLRINH

KOALKLGFQVSQPKKIIIDFSSPNIAKDMIVGHLRSTIIDDSLARIFSVVGHDVLRINH

IGDMATAFGKL/TYLQENPCDYSDLEDLTSLYKKAVVCFTNDEFKKRSQONVALQAKD

ROMATAFGKL/TYLQENPCDJSLESHKAVVCFTNDEFKKRSQONVALQAKD

KOVFHEAFS/PFNVQKSDGSYNYATTDLAAMFYREEDHADKIIIVTDLQSLHFQLLEA

KCVFHEAFS/PFNVQKSDGSYNYATTDLAAMFYREEDHADKIIIVTDLQSLHFQLLEA

KCVFHEAFS/PFNVQKSDGSYNYATTDLAAMFYREEDHADKIIIVTDLQSLHFQLLEA

KUTHEAGSTLAFGHINAIKYSDLSSHRTSDYVFSFEMLRFEGNTAMFLLYAYVRIQGIK

RIDGEROFARVIGINAIKYSDLSSHRTSDYVFSFEMLRFEGNTAMFLLYAYVRIQGIK

RRLGISQLSLEGPPEIQEPAEELLALTLLRFPEALESTIKELCPHFLTDYLYNLTHKFNG

FFRDSHIODSPYAKSRLFLCALAEQVLATGMHLLGLKTLERL 659044 660789

CPn_6571 662179 660749

murA-UDP-N-Acetylglucosamine Transferase

TF#XVNVSSSDFDAKGERRNGIAQVFCCGRLNGEVKVSGAKNAATKLLVASLLSDQKCTL

RNVPDIGDVSLTVELCKSLGAHVSMOKETEVLETYTEEIQCTRVPPTFSNVNRIPILLIG

MLIGRCPESVYVPTVGGDAIGERTLNFHFEGLKQLGVOISSDSSGYYAKAPRGLKGNYIH

ALIGRCPESVYVPTVGGDAIGERTLNFHFEGLKQLGVOISSDSSGYYAKAPRGLKGNYIH

PYPSSVGATENLILAAIHANGRTVIKNVALEALBLIDLVIFLQKAGADITTDNBTIDIFG

TGGLGSVDHTILPDKIEAASFGMAAVVSGGRVFVRNAKQELLTPFLKMLRSICGGFLVSE

SGIEFFGERPLVGGVVLETDVHPGFLTDMQQPFAVLLSQAQGSSVIHETVHENRLGYLMG

LQHMGAECQLFHQCLSTKACRYAIGNFPHSAVIHGATPLWASHLVIPDLRAGFAVVMAAL

LABCGGSIIENTHLLDRGYTNMVGKLRSLGAKIQIFDMEQEELTTSPKSLALRDASL 660749

CPD_0572 662349 664616

CT456 hypothetical protein
IMAAPINOPSTTYOITOTGOTTTTTTVGSLGEHSVTTTGSGAAAGTSGTVTLIADHEMGE
IASQUGSAVSTSAEHSTSTLPPETGSVGATAGSAQSAGLFSLSGRTORRDSEISSSSGS
SISRTSSNASSGETSRAESSPDLGDLDSLSGERAEGAEGEGFGGGLPEST IPHYDPTDK
ASILNFLKNPAVQOKMOTKGGHFVYVDEARSSF IPVRODMSTAES KVSNAKTKENITK
ASILNFLKNPAVQOKMOTKGGHFVYVDEARSSF IPVRODMSTAES KVSNAKTKENITK
RESSGYTPSAARRGAKVETGP IWDDVGGLKGIMKKTTPAPDFSF INETFOGGAHSTSHT
KESSGYTPSAARRGAKVETGP IWDDVGGIKLGITTNVTTEEGGGTNITSTKSTSTDKVS
GPGTPVGATVVPNIVNLGGIKVDLGGIKLGGITTNVTTEEGGGTNITSTKSTSTDKVS
INSTUSGST IEEDTIQFDDPGQGEDDNAIPGTNTPPPPCPPPNLSSSRLTISNASLNQV
LQNVRQHLNTAYDSNGNSVSDLNQDLGGVVKNSENGVRFPYLIPKTTGDTDPSGQATGG
VTDGGGHIRNIIORNTGSTGQSEGATPTPOPTIAKIVTSLRKANVSSSSVLPOPGVATTI
TPQARTASTSTTS IGTGTESTSTTSTGTGTGSVSTQSTGVTPTTTTRSTTSATTTTSS
ASTGTPQARLPSGTRIVATISLVRNAAGRSIVLOGGGRGQSFPIPPSGTGTONMGAQLWA
AASGVVASTLGOVVNQAATAGSQPSSRRSSPTSPRRK

664691 665413 CPn_0573 CYTLUDY/J
yebC family
yebC family
veDMAGHSKWANTKHRKERADHKKGKIFSRIIKELISAVKLGGADPKSNARLRMVIQKAK
VEDMAGHSKWANTKHRKERADHKKGKIFSRIIKELISAVKLGGADPKSNARLRMVIQKAK
ENNIPHENIERHLKKATSAEOKNFEEVTYELYGHGGVGIIVEAMTDNKNRTASDHRIAIN
ENGOSLVEPGSVILYHFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP
KRGGSLVEPGSVILYHFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP
KRGGSLVEPGSVILYHFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP
KRGGSLVEPGSVILYHFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP
KRGGSLVEPGSVILYHFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP

CPn_0574 665978 665394
No robust homolog present in Genebank/EMBL as of 11/7/98
SAERGFRHBIVMVETVLHNFQRYLSKYLYRVFRFPCRKKTFLSSHRVLARPSFPVDYCPG
KIYDLQEIYEELNAQLFQGALRLQIGWFGRKATRKGKSVVLGLFHENEQLIRIHRSLDRQ
EIPRFFMEYLVYHEMVHSVVPREYSLSGRSIFHGKKFKEYEQRFPLYDRAVAWEKANAYL LRGYKKRVGGGYGRA

CPn_0575 666524 665982
Yhhy-amino Group Acetyl Transferase
SIFCRVWRSFMTAEKQNTGILGLEIRYTLPSDATYMLKWLNDPKILRGFPIQTEAEIRET
VNFWVGFYRYHSSLTAVYNSNVAGVATLV/LNFYVKVSHHALISIIVGEEFRNKGIGTALL
NNLIHLAKTRFKLEVLYLEVYEGNPALHLYQRFGFVEVGRQNRFYKDEIGYLAKTIMEKD

CPn_0576 cc7543 666494 prfB-Peptide Chain Release Factor 2 (natural UGA frame-shift) MQENLDKRLEALRTEISLAARSL

CPn_0576.1 667598 prfB-(natural UGA trame-shift) MQENLDKRLEALRTEISLAARCL

CPN_0577 64786 467886 668155

POWITY AND SOUDTEN DISCRETE INKNAMEN TRKHNICODÓKNKUNIT ENHANCE (AMA) COMPTEN DISCRETE INKNAMEN TRKHNICÓDÓKNKUNIT CRITTAN TA COMPTEN DISCRETE INKNAMEN TRKHNICÓDÓKNKUNIT CRITTAN TA COMPTEN DISCRETE IN TRANSPORTATION TO CRITTAN TA COMPTEN TO CRITTAN TA COMPTEN TO CRITTAN TA COMPTEN TA CRITTAN TA COMPTEN TA CRITTAN TA CRI

CP0_0578 668388 OOR 188 GAG 165 CPE_D374

you'l-phosphohydroluse
TTHUSETVLES SLATLE LLAF WASHETEIMMLEPTATEWRLPKKHAILLISCUK LAQUUDLAF
TTHUSETVLES SLATLE LLAF WASHETEIMMLEPTATEWREPKKHAILLISCUK LAQUUDLAF
THIKAVDEKELKINSKISK LAKUSPULTVFYXIDLLCHARLEDKERLEFTEIMTLEADIAVFATL
CHIIDYSUY I SRNTKGETTO LPEEKURH TORAT LAVMQ'ILFSS DYYKYDPHLTTOEPHEDI.

LKLLKNTPLTLLHNTTHVIPNTLNIVGLGDLFARGFH LTRLAQYYMDFVLSGHSHGPQVTLSWPKFARKFFERL NRGLGULKRIRFCSPPEICYLTCSYD NYDPSLPGLLLSHNPDG PYLARG(FVTKEGKOLYV

CPn_0579 669310 669993
ygbP/yacM-Sugar Nucleotide Phosphorylase
KEFASAPLLKSATSHVPMIKSSLILLSCGQCTRFGSKIPKQYLPLNCTPLVLHSLKILSS
KEFASAPLLKSATSHVPMIKSSLILLSCGQCTRFGSKIPKQYLPLNCTPLVLHSLKILSS
LEPOIAEVILVZDPTVOETFOEYPVSFAIRGERRQDSVFSGLQQVSYPW/IIHDGARPFIY
CHEHH LASTAGE FORENAAJETFYTIKOPHTVPMICPDNICATHTACALL
LAFEKYTTVJDJTEAAETFOETDQLVFPKHDQIKLTYPEDLTTACALL

CPn_0580 669936 670793

trua-Pseudouridylate synthase I
ASSNOMFLPRRSNDCPSPPMTKVALLIAYQGTAYSGWQQPNDLSIQEVIESSLKKITKT
ASSNOMFLPRRSNDCPSPPMTKVALLIAYQGTAYSGWQQPNDLSIQEVIESSLKKITKT
RTPLIASGRTDAGVHAYQQVAHFRAPDHFLFANANLTKKALNATLEKDIVIRDVALFDDN
FHARYLAIAKEYRYSLSRLAKPLPWORHFCYYPRHPFSTELMQGGANLLIGTHDFASFAN
HGRDYNSTVRTIYTLDIVDKGDSLSIICRGNGFLYKMVRNLVGALLDVGKGAYPPEHLLD
ILEQKNRREGPSAAPAYGLSLHHVCYSSPYNNFCCEQCSVSTSNEG

CPn_0581 671533 670745
Phosphoglycolate Phosphatase
EGLRWRSYKSFLRQCWIYSMLV5DEFQLCLRSGMYLEDYDVFFFDLDGLLVDTEPCFYRA
FLOACAEFSLEVHMDFSTYYSHTTLGTEIFSKKFIEQYPQAQEYMAEIFAKRLQIYYKSL
EHAGPALMEGVEAF IELVLSLNKTFGVVTNSPRDATHTLRTMYPILNKFLFWTRENYAR
PKPYGDSYDYAYRTFAREGMKVIGFEDSVKGLRALSKIPATLVCINSMAEITPEDYPELK
GKEFFSYPSFDVLTEHESQQKLL

CPn_0582 671305 672177
CT465 hypothetical protoky control of the c

CPn_0583 672349 672717
CT466 hypothetical protein
IVLSFFLGKTKVTPRFLMNERTLLLLLKKKKGLFLAILDLTQTESSLTTPELEKVLKQKK
IFLSCIDRVDLQIKEFRHAFSSELPQDIQEELEEIRDVIIRILDTDKRNYAQKKKEFGIY
ERP

CPM-0584 672659 673798
at65/ntr8-2-Component Sensor
IRTMATHHRKKRNLVFMAVEDSNLHPPAYELLEIKARITQSYKEASAILTAIPDGILLL
SETCHFLICKSOAREILDIDENLEILNRSFTDVLPDTCLGFSIQEALESLKVPKTLRLSL
CKÉSKEKEVELFIRNNEISGYLFTOIRDRSDYKQLENAIERYRNIAELGKMTATLAHEIR
NERGIVGFASILKKEISSPHORMLSSIISGTRSLANILVSSMLETYKSQPLNLKIINLQ
DEESSLIPLLSVSFFNCKFVREGAQPLFRSIDPDRWNSVVMNLVKNAVETGNSPITLTLH
TSGBISVTNRGTIPSEIMDKLFTPFTTKREGNGLGLAEAQKIIRLHGGDIQLKTSDSAV
SEFILIPLLAALPKERAAS

CHIOSES 675880 673865

*SIMILARITY tO CP INCAL2

*SIMILARITY TO CRICAL2

*SIMILARITY TO CRICAL2

ISTERRILARINFSIGDCSSNMATPAOKSPTFODPSFVRELGSNHPVFSPLTLEERGEMA
ISTERRILARINFSIGDCSSNMATPAOKSPTFODPSFVRELGSNHPVFSPLTLEERGEMA
ISTERVALORITY OF THE SIMILARITY OF THE STATE O

CPTED586 675993 677183
acC/ntrc-2-Component Regulator
KEKMNPSRCENMAIKNILVVDOEPLLRDFLSELLTSQGFIPDTAENLRNALQMIRSRDYD
KEKMNPSRCENMAIKNILVVDOEPLLRDFLSELLTSQGFIPDTAENLRNALQMIRSRDYD
KUYISDMSMPDGSGLDLIKI IKQSSPHTPVLVVTAYGSIENAVEAMHQGAFNYLTKPFSSF
ALFAFISKAEELKNLVHENLFLHSQTTPDSHPLIAESKAMKDLLAIAKKAASSSANIFIK
GESQCCKEVLSFFIHHNSPRANHPYIKVNCAAIPETLLESELFGHEKGAFTGATYKKAQR
FELAHKGTLLDEITEVPVNLQAKLLRAIQEKEIEHLGGTKTLSVDVRILATSNRKLKEA
IDDKSFRQDLYYRLNVIPLHLPPLRDRQDDILPLANYFLNKFCRMNNTPLKTLSPKAGEL
LLNYPWPGNIRELSNVLERVVILENTSLLTEDMLALA

CPn_0587 677378 678124

*YVYD_Bs conserved hypothetical protein
SYGELFILSTLIKHHVTLGDKMRPHRKHVSSKSLALKQSASTHVEITTKAFRLSMPLKQL
ILEKSDHLPPMETIRVVLTSHKDKLITETHVVASHGKEILQTKVHNANPYTAVIMAFKKI
RTHANKHSNKRKDRTKHDLGLAAKEERIAIQEEQEDRLSNEWLPVEGLDAWDSIKTLGYV
PASAKKKISKKKMSIRMLSQDEAIRQLESAAENFLIFLNEQEHKIQCIYKKHDSNYVLIE
PSILKPGFGT

CPn_0588 678033 678626
CT469 hypothetical protein
TSKSIKSNAFIKNMTATMSLLNLPSSQDSASEDSTSQSQIFDPIRNRELVSTPEEKVRQR
LLSFLHHKLNYPKKLIIIEKELKTLFPLLMPKGTLIPKRRPDILIITPTTTDAQQNTHN
LGDPKPLLLIECKALAVNQNALKQLLSYNYSIGATCIAMAGKHSQVSALFNPKTQTLDFY
PCLPEYSQLLNYFISLNL

CPB_05HP 678634 679395
CT47/0 hybothetical forcein
SMO(CYTRY) REPLIKHITUTE TERGUPT FARGORT LOCDY RETLYP I GLOKYT
LHIRYCORUPKUTEG RELIKHITUTE LT PEGLET FFARGORT LOCAL GROWK KANSHKUP SLEENF
HIR I PEGCON PEPFAA I FYUKULOY EGILD LT PACCUCKAGUPYACYKYCHIKUCKKHOHK
QALGI FREBED LIQAI I HAKOF GELLAI TAEPTIA TAEKI FYUP DO JOEEKKUERNOGEDP
YHELLELSKVI HY

Cfm_0590 680136 679516 CT471 hypothetical protein LFLYGDHNLGFACRYLFFF GASFUNGLLGVPTWLHEBERFYTHREDESKSYPOME NMETOAORKKRVEFILTGEFALL TLNYQSGFGHLRAKCRUVYPVLYALNFSCSSCKMDM DFRGKWNRSSTITISNOKESINÚKLEKDVGVIVNTKTSLKGNVCPGSTFIKOSWGVWNKI YHNDLVGFSEVTLIFNVSSENGTITFS

CPn_0591 680 64 681020
yage family
SLIMRCTAYCTASAYNLHVLFHLLKPRYFTILSREYVLANLDSTOASNOLAIFFPFGVAV
SLIMRCTAYCTASAYNLHVLFHLLKPRYFTILSREYVLANLDSTOASNOLAIFFPFGVAV
ANG AGGIVENTEPETTIYV FOUR ELEMENT OF A GREEN FOR THE LONG FALAFF PEUPLELAGUNL
HSDILDEPDFFWDHPETO\$IYRDVLSCLDIEARINVLIV

CPn_0592 681132 681461
yidd family
LYSKMFSMSFKRFLOO PVRICLLIIYLYOWLISPLLGSCCRFFPSCSHYAEQALKSHGF
LMCCWLSIKRIGKCGFMHPGGIDMVPKTALQEVLEPYQEIDGDSSHFSE

CPR_0593 682494 681391
CT414 hypothet_cal_protein
VLGAKCMAFKRKTRMLMOVILLEVGLIMILFLLLFYSAIFRKDIYKLHLFSGPLIAKSRK
VYLSEDFINEISOASLDDLISLFKDERYMYGRPIKLMALSVAIASHHIDITPVLSKPLTY
TELKGSSVRWLLBNIDLKDFFVILDYLRCHKYPYTSKGLFLLIEKMVGEGWYDEDCLYHF
CSTPEFLYKTLLVGADVQASSVASLARNVIRCGSERFFHFCNESERTSMISATGRQKVL
KSYLDCESSLAWLLLLVHDSDVVLHFFCDEDLEKVIRLMPQESPYSQNFFSALQHSPRRE
LACMSTQRVEAFRVGEDQDEEYVVQDGDSLWLIAKRFGIPMDKIIQKNGLNHHRLFPGKV
LKLPAKOS

pheT-phenylalanyl trna Synthetase Beta
ntchttovtvkslvktslrlssmripitlldtyfseplstkeileacdhigieaeientt
lysfasyltaxilitiphenaklrvatltdgekehqvvcqaphceaglivallalpgakl
lysfasyltaxilitiphenaklrvatltdgekehqvvcqaphceaglivallalpgakl
fpseqqaytikksklrqvesqomccgadelgldelqiqerallelpeatplgedlatvlg
ntsleisltphlghcasfiglareichvtoanlvipkefsfenlpttallmondpdicpf
fsyvvitgisaqpspiklgeslqalkokpinatudityihlisiqqelhaydashvalds
fsyvvitgisaqspiklgeslqalkokpinatudityihlisiqqelhaydashvalds
fsyvvitgisaqspikltpesspypvvrddhsllgiggvmgakapsfgettittvika
laveklepseltilagetvllpsesayrftrgidponvvpalqaaihyileifpeatispiyss
geicrelkevalrpktlorilgksfsiellsoklosfstropetsllvkypsyrhdin
eeiddveeicrteswaietonpvscytpiyklkretagflanagloeffpdlldpetva
ltrvekeeislogskhttvrssllpgllksaathlnnqapsvqafeigtvyakhdedvq
ltrvekeeislogskhttvrssllpglksaathlnnqapsvqafeigtvyakhdedvq
cogvlrihkosfatlogvheelakkaqikhpvffaelhldlcomlkkttklvkpayp
ssfroltitveedipaallroklhegskwlesvtiisiyodksletrnknvslrlvfod
ybrilsnodieeycrlvallnelltdtrgtins

DED 0595 684943 685926

ET476 hypothetical protein

RDYOFHKOLLFCVCVFAMSCSAYASPRRODPSVMKETFRNNYGIIVSGGEWVKRGSDGTI

RDYOFHKOLLFCVCVFAMSCSAYASPRRODPSVMKETFRNNYGIIVSGGEWVKRGSDGTI

RVVLKMBATLHEVYSGGLLMGEITLTFPHTTALDVVQIYDOGRLVSRKTFFVNGLPSQEE

LFNEDGTFVLTRWPDNNDSDTIRKPYFIETTYGGHVIEGSYTSFNSKYSSSIHMGEGVRS

VFSSNILLSEETFNEGWMXYTTFYPNDEPSITHYGNGOPMGLRLTYLGGGIPNTIEE

WRYGYDOGTTIVFRNGCKTSEIAYVKGVKEGLELRYNEQEIVAEEVSWRNDFLHGERKIY

AGGYOKHEWYYRGRSVSKAKFERLNAAG

CPL_0596 685930 686457
ada-methyltransferase
FAVMADDTLIPKLMYNSLSQACSEGLLIAKYPPLQVIVHFDNNLVVKTHLSVAPVFSCLF
LGPAAHKAMGETVLACSRYANKEHPPFSSHFAKDLIPSQYLEILNCVAEIPFGEQQTYAE
IAKKTDTHPRTVGAACKQNPFLLFFPCHRVVGSHGERNYVLGPVIHEILLKFENSY

CPn_0597 688215 686479

opc-0ligopeptide Permases
MCKHPSFYQRFLSAYKNLLASLSWKFFISVALIGIYAPLFASSKPLLVTMHGEIFFPLL
RYLFFBCYYTKPUDLFFNVLMVTFPFFILSFKLTRGMLRAWLIGLCIISQCMIFAWAYSG
KVODPALAENLKWMRAEKVRENISKNSEMMLLEPKDTRTWEMERRYMSTYEQLGILIKA
KYRKKQEASVKKYQVAFEEKRQSPMFTLRHLEMKNEGICLKRLQORVDKMGPYEMAQQA
WNRATDMYRFFLMALTRIEHELRLADYNNWGOPEDLCIAYANVEKRAEPYKKSLLEIRQV
LEDYAKLRSAISFIQDKRUMIEKESEDLRILINPFSSFHWEDDAGGSREMNKYVPMWQL
SEVTRKDLLAALVFGIRIALVWAGIGITIALAIGIMIGLVSGYFGGTVMLISRFTEIWE
TMPVLFILMLVISITOQKSLLLMTVLLGCFSWTGFSRYVRIEVLKQRDRGYVLAATNLGY
SHYYIMVHQILENAIVPVISLVPFAMMAMISCEAGLIFLGLGEESSASWGNLMREGVTGF
PAESAVLWPPAIILTMLLIAIALIGDGVRDALDPRLQDS

CPn_0598 689712 698219

ODDB-01igopeptide Permease

ECGSVLKYILKRLVLIPLITLTFAIVSINFVILNAAPGDVLEEKSRDALGEAGKSDKMRSY

EKGPDRYLGFREHYGLTLPIFFNTRPKITHKKIOTALDELANANNTTPSAKNAAKSLVYWG

DCAKFVMPALLFEADDASRDDKYRHIAADLFIRGVLOGFVEPNLSPEGRAONKEIAESN

AFLVRQLNEEDLDTKVEALKGWFODHGOTEVPCYSSKOFWKTFFLETRFARYMSRVLRLD

FGTLRNDAHKTVISEVIKRLRCSLVLSILPHIVGFVLCQIFGMIMALKRNRWIDHSLNFI

FLILFSIPVFVAVFWILDNFVINKTIPFTTIPMPYSGLRSPPEVFNELSTLGRIFDLVSH

GFLPFCAVSYGALAAQSBLSRSIFLEVLSQDFICAAKARGLRWFDILYKHVGKNAAVSIV

TSLASSLGTLLGGALVVETLFNIDGFGNFFYQAILNRDHNVVLFSVLVGSALSLVGYLLG

DICYYLLDPRVQLEGRRI

Chn_9600 c02156 c71827 L97827 No robust homelog present in Genelank/EMBL as of 1177798 HGYMKIKKSFOYSLCOAKRFONMLPNHFDPCLOPVNU FSSLLKEETCSLNRAKOHLLYKILRDFNTMOHLRSLO

LAYGEL I ILLSKYQQKT

693092 692736

CT483 hypothetical protein

OFFRIHADDIINSMDEITPNYPLLRQDSLWNRVRVSWRADLSVSSRYEIASAIAILGLLV
AFCASAAVSIIFTANPLAQVFIDOCLALGLLPIPLVIGLLIIGIIVLLYGIYLFPQQRE

64417 WHA 643104 John 602

TITIAN HYPOTHAN HALL PLOTHIN GASTON GASTON TO THAN HYPOTHAN HYPOTH

694136 695185 CPH_0603 694136 695185
hemZ-Ferrochetalase
wKiMRIVLMQCLVSLFLAKKVTVTTPAYLLANFCGPRHAKDLQEFLISLLTDRDVTGTF
LPRVLHRHLFTFIAKKRVPKVLPOYQSLQNMSPIYFDTETLAKTLSEILRAPVIPFHRYL
PSTHEKTLLALRTLHTRHVIGIPLFPHFTYSVTGSIVRFFMKHVPEIPISWIPQFGSDSK
FVSLITCHIRDFLOKLGILEKECCFLFSVHGLPVRYISQGDPYSKQCYESFSAITTNFKQ
SENFLCFQSKFGPGKWLSPSTAQLCQNIDTDKPNVIVVPFGFISDHLETLYEIERDYLPL
LRSRGYRALRIPAIYSSPLWVSTLVDIVKENSTVVAEELIKSGKKHTGIR

CPn_0604 695981 695196
fliy-Glutamine Binding Protein
CKKRQNSEAQLNVKIKFSWKVMFLICLLAVGLIFFGCSRVKREVLVGRDATWFPKQFGIY
TSDTNAFLNDLVSEINYKENLNINIVNODWYHLFENLDDKKTQGAFTSVLPTLEMLEHYQ
FSDPILLTGPVLVVAQDSPYQSIEDLKGRLIGVYKFDSSVLVAQNIPDAVISLYQHVPIA
LEALTSNCYDALLAPVIEVTALIETAYKGRLKIISKPLNADGLRLAILKGTNGDLLEGFN
AGLVKTRRSGKYDAIKQRYRLP

696737 696150

CPI_0003
yhbF-Methylase
LRKLCSSRGDVRILAGKYKGKSLKTFSNPHIRPTSGLVKEAFFSICREDIEGAAFLDLFA
GMGAIGFEALSRGAASVVFVDISIKAIQLIHTNSALLGEQLPVVIFRQDAQSAIQRLIKQ
KRSFDLIYIDPPYELCNCYVETLLQKIVSGNILNPEGTLFLENASDEEIACEGLTLRRRR

CPn-0606

CPH-0606 697492 696707
CT#88 hypothetical protein
SSYSRRQLRYTGSLOMHIYGLADLHLALGVPEKTMEVFGDPWIGYHOKICSEWQAVVHP
EDITVLLEGDISWAMMLSEAHKOFAFIGOLPGTKYMIRGNHDYWSSASTSKILQALPPSLY
YLWGGFALLTPHLAVVGVRLWDSPTICVKKENFLTPSTQEQSYTEQDEKIFLRELGRLKR ARALPKEVTEVIVMTHYPPISSDGTPGPISEFLEADGRVSLCLFGHIHKVQRPIDGFGN IRGIHYILVAADYVNFVPQEVM

CR__0607 698910 697573
gf@C-Glucose-1-P Adenyltransferase
NRBIOMIENDFPEASNFESSHFYRDKVGVIILCGEGGKRLSPLTNCRCKPTVSFGGRYKL
IDIFISHAISAGFSKIFVIGOYLTYTLQOHLFKTYFYHGVLQDQIHLLAPEARGGDQIWY
OGFADAIRKNLLYFEDTEIEYFLILSGOQLYMMDFRSIVDTAIRTHVDMVLVAQPIPEKD
AYRKVLDIDSEGKLIDFYEKPQEKEVLKRFQLSSEDRRIHKLTEDSGDFLGSMGIYLFR
RDSLFSLLREEGNDFPGKHLIOAQWKRGQVOTLLYMGYWADIGTIESYYEANIALTOKPH
AEKRGLNCYDDNGMIYSKNHHLPGAIITDSMISSSLLCEGCVINTSHVSRSVLGIRSKIG
ENSWYDQSIIMGNARYGSPSMPSLGIGKDCEIRKAIIDENCCIGNGVKLQNLKGYIRYDS
PÖKKLFYRDNIIIVPGATHIPDNYIF POKKLEVRONIIIVPOGTHIPDNYIF

CED_0608 699690 699016

*\lightarrow{\text{VFTdine}} 5'-Monophosphate Synthase (Ump Synthase)-truncated?

VSFLYFYKNGRRLWRMANYLOBAKLRGQAVALLYQIGAIKFGKHILASGEETPLYYDMRLV

ISSFEVLOTVATLIWRLRPSFNSSLLGGVPYTALTATSISLKYNIPMVLRKELQNVDP

SDATKVEGLFTPGQTCLVINDMVSSGKSIIETAVALEENGLVVPRALVFLDRRKEACQPL

GPGGIKVSSVFTVPTLIKALIAYGKLSSGDLTLANKISEILEIES

699986

CPH_0009 0990/2 0990/2 0990/2 07900 CT490 hypothetical protein
QNTKNSLIRENMLIRLFLGISLPKGFPLYLEPPLVLATFQGTQFVGTYSEATNPLYYONL
NLNYHYTQELLYKAVPCNYKSIYREIPLIIFPEVLIGSTPTQSTE

CPn_0610 701450 700029 rho-Transcription Termination Factor rho-Transcription Termination Factor
RIFILFROSIMKEERSSELLPRVKETKKHAYVSMOEKSCVGECAVVASESEEARSVTVTK
IAKLQRMGIEELNILARQYGVKNIGSLTKSQVVFEIVKAKSERPDELLIGEGVLEVLPDG
FGFLRSFTYNYLPSAEDITVSPAQIRRFDLKKGDTIIOTIRSPKEKEKYFALLKVDKING
STPDKAKERVLFENLTPLYPNORIVMEMGKDHLAERVLDLTAPIGKGQRGL VAPPRSGK
TVILQSIAHAIAVNNPDIVLIVLLIDERPEEVTDMIRQVRGEVVASTTDEQPERHIQVAE
MVIEKARRLVEHGNDVVILLDSITRLARAYNTVQPHSGKILTGGVDASALHKPKRFFGAA
RNIEGGGSLTILATALIDTGSRMDEVIFFEFKGTONMELVLDRRLSDRRTVPAIDLIKSG
TRKEELLYHPSELERVYLFRQAIADLTTIDAMHLLLGRLKKTNSNAEFLLSLKE

CPn_0611 702133 701420
yacE-predicted phosphatase/kinase
RRNRRDAKTSEREDGISYDFIRSYSCEYLNWKKLGRMLKLLKVSITTDLSSGKTEACOVF
OELGAYVVSADEISHSFLIPHTRIGRRVIDLLGSDVVVDGAFDAQA/IAAKVFYNSVLLQG
LEAILHFEVCRIIEECYHOSIODGNYPLFVAEVPLLYEIHYAKWFØSVILWMANEDIRRE
RFMKKTGRSSEDFDQRCSRFLNVEEKLAQADVVVENNGTKKELHYKIEEYFYALKGAL

704688 702022 polA-DNA Polymerase I pola-jona. Polymetase t RCHITGLIGTVERPRREYAMKKLFVLDASGFTFRAYFALPEMKNIQOATGAVFGFTRSL NKLTKEESPEYM ISVEDGINNKQSRQATYADYKSMEQKKEEBTPOTALVKEYYSLTGLA YLERESVEADDV TASTAKKAREENYKVYVCTADKDLIGLVVDHVVANMPWADGGVVGTGE VIERYGTTPEMI IPDVLALVGDSODN FPGLPGCQPKKAAAL KOPGSVEGLLENLDAVKGL SOFMLGEROETEKLISKILALLDON I PTEV FTESLTFPDH VVDEEKLTHEYTGCGFRTLVP SKOTEAATYDVO I TKOAESILTIN ILINLVQSIGDTAFAVAYTININLIGIKLEGILALTIGGSOVF FTALEEGITRI KLIFTEKDHELIKERINGPGVENTIKRDCHALLMAST VIRBETSVDLALAEHLTIN SKOTKTGPGGLLVMRGFTETAHRFAKEWINSGLFTGRLVFOPFGGFVAVLPTTKDATL EETHPRIIGHTLEGIT EMPLEKVLFSMERASVPLDVEFJATLEALFETELAVLTEETYDLS CREFNIKSPYCLSDILYNE LOKAKSTPAEVLEALRSEHPIIEKLLS
TYVKALPKOVDSHTORIHE KOAVTKRLACRDENLONIFIRSEEGILLRKAFRISEK
NSYFLSADYSQIELRFLAHLSODKSLKFAFESSEDIHAFTASOVFHVPLEOVSKEORNOA
KTVNFGIVYGOOAFGLAVVLKISIGEXQELIQAYFSRYPEIAHFVEETIGOAAKDLRVTT
MIGRERIIDSWNEFFGSPAASGRFAVNTRIOSSAAELIKLAMLDISQAIKOOMKSRMLL
QIHDELLFEVPEEEIEEMQRLVREKJESAMTLSVPIVVNILIGKNWAEC

704653 CPn 0613 VVMKTLWHE''/ ALLIA COLON GRATIFIZCES DAROGENINDES

KTAPIIAVIEMKDVIASSKNTAKTIONILEUFEKAPLKDRVKGIVIDMDCPGGEVFEIDR KTAPITAVIEMRUVIASSANIAATIQNIEDEFRANDENVANTYOTIGENSGPFFWK ISSMLRWERKGFPTYI/VNYLCASGGYYVSCAATKIYATSSSLIGSIGVRSGPFFWK EGLARYGVESDLLTACKDRAPKNPYTPMTSHDREERQATLDFLYQGFVDIVTQNRPLLTK EKLVHTLGARIFSPEKAKGGEVIDVGATKEGVLDDIVAVCKIEDNYRVIGSGGDGWWKR VASAAASSELVTGMIKHDIYPLSHDAAYIPPYLAL

CPn_0614 707435 705793
adt-ADP/ATP Translocase
VFIRKVCKEFMQSSEVKPFSRLRAYLCPIYKSEFSKFVPLFLLAFFVGFNYCLLKNMKD
TLJVIVGSDAGAEVIPFLKVMGIVPGAVIVTMVYGMLGSRYPROTVFYCFMAAFLGFFFLF
AVIIYPVGDSLHLNSLADKLQELLPQGLRGFIVMVRYWSYSIYYVMSELWSSVVLSMLFW
GLANQITTITEAGREFALINTGLNLSSICAGEISYMMCKTFVAVSFACDSWHSVMLNLT
HLITCSGLIMIMLVYRHIHLTIDTSIPPSRRVLAEEGAATANLKEKKRFKARNLFLHL
IQSRYLLGLAIIVLSYNLVIHLFEVVWKDQVSQIYSSHVEFNGYMSRITTLIGVVSVLAA
VLLTGQCIRKWGWTVGALVTPLVMLVSGLLFFGTIFAAKRDISIFGGVLGMTPLALAAWT
GGMQNVLSGRTKFTFDOTKEMAFIPLSPEDKNHGKAAIDGVVSRIGKSGGSLIYQGLLV
IFSSVAASLNVIALVLLIIMVVWIAVVAYIGKEYYSRAADAVATLKQPKEPSSSIVREAQ
ESVEQEEMAVL 107435

CPn_0615 / 708149 707634
pgsA-Glycerol-3-P Phosphatidyltransferase
LAXIMROFOKILLSLSRLWLLLYFCQEKLHIRLLAIVGAMLSDVLDGYLARRYKATSRLGS
ILDPITDKYFVFVCITVLYMEGSLSIAHLFFICARDLFLIIFVCYLSLVKGWKGYDYGSL
FWGKIFTVVOFIILLGVTAGGEIFWTGLVPLVALGFLYFLERIMDYKKOFLR

CPn_06/6 708704 710137
dnaB-Replicative DNA Helicase
TLTNYESSLLMDKSTGVPLPSPPHSKESEMIVLGCMLTGVHYLNLAANOLYEEDFYYLEH
KIIFAVLQDAFKQDKPIDVHLAGEELKRHNQITVIGGPSYLITLAEFAGTAAYLEEYVDI
IRSWSILRKMISTAKEIEKRALEOPKNVAEALDEAQNSFFKISQSTSVSQYTLVADKLRG
LTTATDKPYLVQLOERQELFLQNAQGDMXSFFTGIPTHFIDLDQLIHGFSPSNLMILAAR
PAWGKTALALNIAENLCFQNRLPIGIFSLEMTVDQLIHRMICSRSEVDSKKISIGDLSGH
DEFQRIVSVINEMQEHTLLIDDQPGLKVSDLRARARRMKESYDIQFLIIDYLQLLSGSGTL
RATESRQTEISEISRMLKTLARELNIPILCLSQLSRKVEDRANHRPMMSDLRESGSIEQD
SDLVMFLLRREYYDPNDKPGTAELIIAKNRHGSIGSVPLVFEKELARFRNYSAFECIS

710481 712316 CPT_U017 /10481 /12316
gidA-FAD-dependent oxidoreductase
LMWTHPIAYDVIVVGAGHAGCEAAYCSAKMGVSVLMLTSNLDTIAKLSCNPAVGGIGKGH
IVREIDALGGIMAEVTDOSGIGFRILNOTKGPAVRAPRAQVDKQLYHHMKRLLENTPGL
HIMQATVESLLDKEGVISGVTKEGMMFSGKTVUXLSSGTFMGLHIGDRNFSGGRLGDP
SOGLSEDLKKRGFPISRLKTGTPPRLLASSINFSCMEEQPGDLGVGFVHRTEPFOPPLP SOGLSEDLKKRGFPISRLKTGTPPRLLASSINFSCMEEQPGDLGVGFVHRTEPFQPPLP QLSCFITHTMEKTKAIISANLHRSALYGGCIEGVGPRYCPSIEDKIVKFSDKERHHVFLE PEGLHTDEIYANGLSTSMPFDVQVDMIRSVLGLENAIITRPAYAIEYDYIHGNVIHPTLE SKLIEGLFLCGQINGTTGYEEAAAQGLIAGINAVNKVFNRPPFIPSRQESYIGVMLDDLT TQILDEPYNFTGRAEHRLLLRQDNACARLSHYGYELGLLSEERYSLVKKQNQLLEEEKV RLQKTFRQYCQSVVSLAKALSPRPEVSYDMLREAFPNDIRDLGAVLNASLEMEIKYSGYID RQKILIGSLEKAESLLIPEDLDYKQITALSLEAQEKLAKFTPRTLGSASRISGIASADIQ VLMIALKKHAHH

712300 713010 CPT_0018 /12300 /13010

[pla-Lipoate-Protein Ligase A
KNMPTTNCIFLDLRGHSILHQLOIEEALLRVANQNFCIINSGAKDSIVLGISRNLNQDVH
ISRAQADHIPIIRRYSGGTVFIDSNTLMVSWIMNSSEASAQPQELLAWTYGIYSPLLPN
TFSIRENDYVLGHKKIGGNAQVIQRHRWHHTTFLWDIDLDKLSYYLPIPQQQPTYRNQR
SHEEFLTTLRPWFPSRDDFLERIKASGSLLFTWEEFLDNELEEILAQPHRKATTVLN

713462 CFI_0017
/1002
/10017
/1002
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/10017
/1 HGSDTLENAAVEIAYFFSKIEVVNASKPLV

714145 CPn_0620 713519 CPT_0620 714145 713519
ruva-Holliday Junction Helicase
DKMYDYIRGTLTYVHTGAIVIECQGIGYHIAITERWAIECIRALHQDFLVFTHVIFRETE
HLLYGFHSREERECFRILISFSGIGPKLALAILNALPLKVLCSVVRSEDIRALASVSGIG
KKTAEKLMVELKQKLPDLLPLDSRVETSOTHTTSSCLEEGIQALAALGYSKIAAERMIAE
AIKDLPEGSSLTDILPIALKKNFSGVNKD

CPn_0621 . 714707 714144 ruvC-Crossover Junction Endonuclease LGRLGSSFKDNKFKVFQESIVSELIIGVDFOTIVAGYAIIAVEQRYQLRPYSYGAIRLSS DMPLPMRYKTLFEQLSGVLDDTQPNAMVLETQFVNKNPQSTMKLAMARGIVLLAAAQRDI LIFEYAPNVAKKAVVGKGHASKRQVQVMVSKILNVPEVLHPSNEDIADAFALAICHTHVA

CPD_0622 715761 714793
CT503 hypotherical protein
RYSVPELDITEKEHLERIERISSGESPHYTHOGREGMEHLLCRWKDADIMEWOOICNITESCY
CORMOGRIM-LOKETODSAINGEHERIHLDYREGLIGALEEEYRRREEAKHODLEKLQOENT
WLOMBLAEKLQOIRHQUBI LDEIKKELLDOVQRTEISBEGRECYBHIKKOLEEDLORYVS
OHDAHGI ELEEEDKOSAAYAEINPLEKKOLIDLQOENDIYIKTYHBEIAKLBEKLOROGGAQ
TESEPROSIEKUTEVOTDLAEKKYAIALLODIVEDOYCOLROLHKEKCMAMPONTKLOHLK
GLUSKEPESSEVOVYPSESKSIZO

717011 CPE_9623 *** CPE_9624 *** CPE_963 *** CPE_

RELWEE:QRNEIVEOKKFOLLPPPAKLISEVIGOTVVI

ADLNESLQALVRESSDL

CPn_0624 718018 717011
gapA-Glyderaldehyde-3-P Dehyrogenase
AMKVYINGFGRIGRLVIRGILKRNSSVEVLAINDLVPGDALTYLFKFDSTHGRFPEDVRC
EADHLIVGKRIGFLSERNYQNLPWKDLGVDLVIECTGLFTKKEDAEKHIQAGAKRVLIS
AMKGDIGTFFVMGVMHKTENPEKDFVIGNASCTTMCLAPIAKVLLDNFGITEGLMTTVHA
ACKTRALZIJE ELEPHIMRVIH GELDHI HAADENAVANTLUDEELKERVLTTMAFRNYLEES/ PRVEDTYREDFUTTYDD CHAMKQAUETDEKG EENTEDEQVVOODF IGGEYGS IFDALAG IALNDRFFKLVAWYDNETGYATRIVDLLEYVEKNSK

CPn_0625 718488 718060
rll7-cl7 Ribosomal Protein
VMQHARKKFRVGRTSSHNRCMLANMLKSLIHYERIETTLPKAKELRRHADKMITLAKKNS
LAARRIAIGRLMYRYNKLTSKEARQAKGGDTSVYNVDRLVVNKLFDLGNRFVERKGGYT RILKLONRIGDNAOKCIIEFLAS

CPn 0626 719670 718495

CPH_0626 /196/0 /18495

TPOA-RNA POLYMETASE Alpha
WLPAKKKAQSVVLGKEKGMSDNAHNLLYDKFELPEAVKMLPVEGLPIDKHARFIAEPLER
CMCHTLGNALRRALLIGLEAPAIISFAMTGVLHEYMAIEGVIEDVTNIILALKKGALLKKY
PMODSSLGRTTYOLKASISIDASDLAAMAGGKEVTLODLLQEGDFEAVNPDQVIFTVTOP
IQLEVVLRIAFGRGYTPSERIVLEDKGVYEIVLDAAFSPVTLVNYFVEDTRVGQDTDFDR LVLIVETDGRVTPKEALAFSTQILTKHFSIFENMDEKKIVFEEAISIEKENKDDILHKLI LGINEIELSVRSTNCLSNANIETIGELVIMPEPRLLQFRNFGKKSLCEIKNKLKEMKLEL GMDLTQFGVGLDNVKEKMKWYAEK IRAKNTKG

CPn 0627 720059 719640

rsll-Sll Ribosomal Protein FLIRSRVLVKNQAQAKKSVKRKQLKNIPSGVVHVKATFNNTIVSITDPAGNVISWASAGK VGYSGSRKSSAFAATVAAQDAAKTAMNSGLKEVEVCLKGTGAGRESAVRALISAGLVVSV IRDETPVPHNGCRPRKRRV

CPn_0628 720461 720063
rsi3-si3 Ribosomal Protein
DAYTILREAGRMPRIIGIDIPAKKLKISLTYIYGIGSARSDEIIKKLKLDPEARASELT
EEEVGRLNSLLQSEYTVEGDLRRRVQSDIKRLIAIHSYRGQRHRLSLFVRGQRTKTNSRT RKGKRKTVAGKKK

CM 0629 721881 720487

sect - Translocase KIREFRPYMTTLROFFLITELROKLFYTFALLTACRVGVFIPVPGINGELAVAYFKOLLG KIĘĘFRPYMTTLROFFLITELROKLFYTFALLTACRVGYFIPYPGINGELAVAYFKOLLG
SCÓULFQLADIFSGGAFAOMTVIALGVVPYISASIIVQLFLVFMPALOREMRESSDGKR
REGRLTRLFTVALAVIQSLLFAKFALRMILTIFGIVLPTLLSKLFGVPWIFYITTVVVM
TTGTLLLMWIGEQISDKGIGMGISLIIALGILSSFPSVLGSIVNKLNLGSQDSSDLGLIS
IBĖBALVFYFVLITTILLIEGVYKLIPVYJRARVIGREVPUGGGSYLPLKVMYAGVIPVIF
ASŞELMFPATIGGFIASSSMWKRIAALLAPGSLVYSICYVLLIIFFTYFMTATOFHEPG
IRŠEMKKNNAFIPGIRQGKPTQHYLEYTMNRVTLLGALFLAAIAILPSLLGCLLRVDSNV
SYELGGTAMLIVVGVVLDTMKQVDAFLLMRRYDSVLKTDRTKGRH

CER 10630 722316 721885
rFIS-L15 Ribosomal Protein
MIKEESLFDISERKRRKKLLGRGPSSGHGKTSGRGHKGDGSRSGYKRRFGYEGGGVPLYR
RVPTRGFSHKRFDKCVEETTTGHAELFQEGEAITLDALKAKKAIARQAVRVKVILKGDL **EXTFVWODTAVVLSOGVONLLGIT**

CEn_0631 722812 722312

TSJ:SS Ribosomal Protein

REMISLSKNSHKEPOLEEKVLVVNRCSKVVKGGRKFSFSALILVGDCKGRLGYGFAKAMEL

TDAJRKGGEAAKKNLHKIEALEDGSIPHEVLVHHDGAQLLLKPAKPGTGIVAGSRIPLIL

ENGIKDIVAKSFGSNNPMNQVKAAFKALTGLSPRKDLLKRGAAIND

CPH_0632 723354 722827
rtlfe-L18 Ribosomal Protein
KGTTSSWLVNLLQVFAPNVLLMLIKVREFVMKMMMSVVKLVKLRKQAKNRSRVMESSLCK
KSERKRRRALRVRKVLKGSPTKPRLSVVKTNKHIVVQLIDDSIGKTLASVSTLSKLNKSQ
GLTKKNQEVAKVLGTQIAELGKNLQLDRVVFDRGPFKYHGIVSMVADGARZGGLQF

723760 723209

CPH_0633 723700 723209
r16-L6 Ribosomal Protein
SMSRKAREPILLPQGVEVSIQDDK:IVKGPKGSLTQKSVKEVEITLKONSIFVHAAPHVV
DRPSCMGGLYWALISMMVQGVHLGFEKRLEMIGVGFRASVQGAFLDLSIGVSHPTKIPIP
STLQVSVEKNTLISVKGLDKQLVGEFAASIRAKRPPEPYKGKGIRYENEYVRRKAGKAAK

CPn_0634 724215 723787
rs8-S8 Ribosomal Protein
ESSIKRKRIYMGMTSDSIADLLTRIRNALMAEHLYVDVEHSKMREAIVKILKHKGFVAHY
LYKEENKRAMRVFLQYSDDRKPVIHQLKRVSKPSRRVYVSAAKIPYVFGNMGISVLSTS
QGVMEGSLARSKNIGGELLCLW

724763 724206

CTIS-L5 RIDOSOMA1 PROTEIN
GERKANMSRLKKEYTEBIRKSLFEKFGYANKMOIPVLKKIVLSMGLABAAKDKNLFQAHL
BELTMISGOKPLUTKARNSIAGFKLRECOGIGAKVYLRGIRMYDFMDRFCNIVSPRIRDF
RGFSNKGDGRGCYSVGLDDOOIFPEINLDRVRRTYGLHITMVTTAQTDDECTTLLELIMGL

r124-L24 Ribosomal Protein

FKEKEVMKKON I RVGDKVF I LAGNDKGKEG VLSLTEDKVVVEGVNVRIKNIKRSQQNPK CKRISTEAPTHISMVRLTTAGEPAKLSVKYTEQGRELWGRRPDCTSQLYRLVRGKKG

7114-1.14 Ribodomal Protein
TERMIOUESULKVADNTGAKKUKERKULGGGRRRYATUGDVIVESURRUEENGGIKKGDV
TPAVIURTRHITTERISSSTLKEDTINGCVIIDDKGNEKGTRIEGEVAREIRDRGFIKISSL

Clm_0n 08 725490 NKKEKVKSMASEPPCSRKVK VSAKMEKTYVRVER I FSHPQYLKVZRSSKKYYAHT ELKVSECOKVK I GETRPLSKLKRWRV I EHVOVÝS

725979

CPn_0639 725979 725743
rl29-L29 Ribosomal Protein
ASGKGINMAAKKDLLTQLRGKSDDDLDAYVHENKKALFALRAENLLQNKVVKVHMFSTHK KNIARALTVKQEPYGKVHG

7116-L16 RIDOSOMAL PROTEIN 1216-L16 RIDOSOMAL PROTEIN 1116-L16 RIDOSOMAL RIDOS

CPL_0641 727092 726409
rs3-s3 Ribosomal Proyein
kgRRIMOGKCCPIGFRTGVYMKWRSLWYGNKQEFGKFLIEDVRIRGFLRKKPSCQGAAGF
VVRMSGKIEVTIQTARFGLVIGKKGAEVDLLKEELRALTGKEVWLEIAEIKRPELNAKL
VADNIARQIERRVSFRRAMKAMQSVMDAGAVGVKIQVSGRLAGAEIARSEWYKNGRVPL
HTLRADIDYATACAETTYGIIGIKVWINLGENSSSTTPNNPAAPSAAA

CPn_0642 727440 727096 r122-L22 Ribosomal Protein RRHSMFRATARYIRVOPRKARLAAGLMRNLSVQEAEEQLGFSQLKAGRCLKKVLNSAVAN AELHENIKRENLSVTEVRVDAGPVYKRSKSKSRGGRSPILKRTSHLTVIVGEKER

727725

TS19-S19 Ribosomal Protein
EIRIMGRSLBKGPFVDHHLLKKVRAMNIEEKKTPIKTWSRRSMITPEMIGHTFEVHNGKK
FLTVFVSETAVGHKLGEFSPTRIFKSHPVKKG

728594

CPI_U648 72594 72772
rilosomal Protein
FIREINSMFKKFKPVTPGTRQLVLPAFDELTTRGELRGTKSKRSLRPNKKLSFFKKSSG
RDNLGHISCRHRGGGAKQLYRVVDFKRNKDGITAKVVTVEYDPNRSAYIALLSYEDGEKR
YILAPKGIQRGDVVVSGGSFFKPGCCTHILKSIFJGLSVNID EDMPSSGGKLVRSAGLA
QVIAKSPGYVTLKMPSGEFRMLNEGCRATIGEVSNADHNLRVDGKAGRRRWMGVRPTVRG TAMINPVDHPHGGGEGRHNGYIPRTPWGKVTKGLKTRDKNKSNKWIVKDRRK

728933

YHLOUGH YHLOUGH CHIZO-LL23 RIDOSOMAl Protein DMKDPYDVIKRHYVTEKAKMLEHLSAGTGEGKKKGSFCKDPKFVFIVSHDATKPLIAQAL EAIYVDKNVKVKSVNTINVKPQPARMFRGRRKGKTSGFKKAIVTFYQGHSVG

729636 728950 rl4-L4 Ribosomal Protein

YREDLMYLLSKFDFSGNKIGEVEVADSLFADEGDGLQLIKDYIVAIRANKROWSACTRNR SEVSHSTKKPFKQKGTGNARQGCLASPQFRGGGIVFGPKPKFNQHVRINRKERKAAIRLL LAQKIQTNKLTVVDDTVFVDALTAPKTQSALRFLKDCNVECRSILFIDHLDHVEKNENLR LSLRNLTAVKGFVYGININGYDLASAHNIVISKKALQELVERLVSETKD

730490 729657

T13-L3 Ribosomal Protein

**T13-L3 Ribosomal Protein

**YLEYFSYCKNLPPLITCFFIERENFLFFLENSISKILSRFVSLFLQEESKSLLLMDKFM

**SHISYMCKKEGMHIFDKDGSLVACSVIRVEPNVVTQIKTKESDGYFSLQIGAEEMNAP

AHTITKRVSKPKLGHLRKAGGRVFRFLKEVRGSEEALNGVSLGDAFGLEVFEDVSSVDVR GISKGKGFQGVMKKFGFRGGPGSHGSGFHRHAGSIGMRSTPGRCFPGSKRPSHMGAENVT VKNLEVIKVDLEKKVLLVKGAIPGARGSIVIVKHSSRT

731636

CPn_0648 731636 730605
CT529 hypothetical protein
FFFKKPCKEVMATNAIRSAGSAASKMLLPVAKEPAAVSSFAQKGIYCIQQFFTNPGNKL
AKFVGATKSLDKCFKLSKAVSDCVVGSLEEAGCTGDALTSARNAQGMLKTTREVVALANV
LNGAVPSIVNSTORCYQYTROAFELGSKTKERKTPGEYSKMLLTRGDYLLAASREACTAV
GATTYSATFGVLRPLMLTNKLTAKPFLDKATVGNFGTAVAGIHTINHMAGVAGAVGGIAL
EQKLFKRAKESLYNERCALENQQSQLSGDVILSAERALRKEHVATLKRNVLTLLEKALEL
VVDGVKLIPLPITVACSAAISGALTAASAGIGLYSIWQKTKSGK

732672 731710

CPH_0649 732672 731710
fmt-Methionyl tRNA Formyltransferase
LNLKVVYFGTPTFAATVLQDLLHHKIQITAVVTRVDKPQKRSAQLIPSPVKTIALTHGLP
LLQPSKASDPQFIEELRAFNADVFIVVAYGAILRQIVLDIPRYGCYNLHAGLLPAYRGAA
PIORCIMEGATESGNTVIRMDAGMDTGDMANITRVPIGPDMTSGELADALASQGAEVLIK
TLQQIESGQLQLVSQDAALATIAPKLSKEEGQVPWDKPAKEAYAHIRGVTPAPGAWTLFS
FSEKAPKRLMIRKASLLAEAGRYGAPGTVVVTDRQELAIACSEGAICLHEVQVEGKGSTN
SKSFLNGYPAKKLKIVFTLNN

CPn_0650 733513 732665

lpxA-Acyl-carrier UDP-GlcNac O-Acyltransferase
SRRNMASIHPTAIIEPGAKIGKDVVIEPYVVIKATVTLCDNVVVKSYAYIDGNTTIGKGT
TIWPSAMIGNKPQDLKYQGEKTYVTIGENCEIREFAIITSSTFECTTVSIGNNCLIMPWA
HVAHNCTIGNNVVLSNHAQLAGHVQVGDYAILGGNVGVHQFVRIGAHAMYGALSGIRRDV
PPYTIGSGNPYQLAĞINKVGLQRRQVPFATRLALIKAFKKIYRADGCFFESLEETLEEYG DIPEVKNFIEFCQSPSKRGIERSIDKQALEEESADKEGVLIES

733975 733517 CPn: 0651 LaDZ-Myristoyl-Acyl Carrier Dehydratase
MNQ65VIKHRELLDLLPHRY9FLLVDKVLSYDIEARSITAQKNVTINEPFFMGHFPNAPI
MFGVLIELALAQAAJVLIGULEJBRNKRIALFLGIQKAKFRQAVRFGDVLTLQADFSLI
SSKCKKAWAQARVDSQLVTEAELSFALVDKESI

CPn 9652 734880

CPG_0652 7 1480 733990
LpxC Myrisroyl Glong Deadetylase
KRNSI (GDSCHAPFYMLERFORTLKREVRYSC/VG HLD/KSSTLHLOPACTNTY)IVFORGS
ASCAMFENVPALLICHVYTTMR:TPTLSRC/SAVIATVEHLMAALRSNNI DNL HOCGEER PI
GDG/SSNAVELLICVAG (GEQEDAVS LARLTPP/YYOHGDIFLAAFPS/DELK LS/YTLHYPO
SCTIG/TOYKS/LVINEESPROEIA/CRITFAL/MELCFLMEKGLIGGGCLDNAVVFKDEGII
GROGJGRADEPVRHKILDI.UGDESLAGREPVAHVLAVGGGHSSNIAFGKKILEALEL

CPr. 0653 736499 734868 CULE-Apolipoprotein N-Acetyltransferd
GEPVLRIFCFVIGWCLIAFAOPDLSGFVSILOAACCY
SLEPLKKPSLPLRTLFVS
CFFWIFTIEGIHFSWMLSDQYIGKLIYLVWLTLITILSVEFSGFSCLLVAIVRQKRTAFL
WSLECWWA IEMLRFYGIFSCMSFDYLGWPMTASAYGROFGGFLGWAQQSFAVIAVNMSF
YCLLLKKPHAKMLWVLTLLLPYTFGAIHYEYYKHAFQQDKRALRVAVVQPAHPPIRKLK
SPIVVWEQLLQLVGPIQQPIDLLIFPEVVVPFGKHRQVYPYESCAHLLSSFAPLPEGKAF
LGNGDCATALSQHFQCFVI IGLERWYKKERNVLYWYNSAEVISHKGISVGYDRRILVPGGE
VID YCKFGCLLYPDJEFPYYALGCKRLPGGRSGFYDVRGRDPRGGTTTCTGEFFGYRLGVS
ACAGEGCTHLTDICTGVFGGGGGKVHFHIEMLPRAGGGMPTYRACOTTTTAAVDGLGPILK
TLIFYDPWETKARGGYTLETDLFLFNYKTGYGYGGTFMILIAFGAVGYLSXGFLGYRLLAK
KEIR

CPn_0654 737051 736503
vdlD/ycia-acyl-coa Thioesterase
KKIIDFLSVDRYYRNQEYPIKILSVESTMLKKKPVSFSCIDGHIYKIFPNDLNANNTVFG
GLLMSLLDRLALVVAERHTESVCVTAFVDALRFYAPAYMGENLICKAAVNRTWRTSLEVG
VKWAENIYKQERRHITSAYFTFVAVNEDNQPIPVHQIVPETPEEKRRYNEADRRRQARL

737856 737101

CPH_0655 737856 737101
dnaO-DNA POL III Epsilon Chain
KEIMSLIKDTVFTCLDCEMTGLDVKKDRIIEIAAVRFTFDSVIŠSIEFLINPERVVSAES
ORVHHISNAMLROPKIAEVFPOIKAFFKEGDYIVGHSVGFDLQVLAOEMERIGETFLSK
YTIIDTLELAKEYGDSPNSLESLAVHFNVPYDGNHRAMKDVEININIFKHLCKRFRTLE
OLKQVLAKPIKMYMPLGKHKGRCFSEIPLAYLQWASKMDFDSDLLFSIRHEIKHRQKGT

CPn_0656 737842 738048
No robust homolog present in Genebank/EMBL as of 11/7/98
THNFLLLPLSLFDILLTVEGFLCLTLYFASVQRMPCEQKRVPGNLYYYYIAAHSSLCLSV

738476 738051

YjeE (ATPase or Kinase)
PMGRYRRVSHSSQETLLLGTELGQVLVPGAVLLLFGDYGAGKTEFVRGIVSGYLGDTIAE
EVASPSFSILHVYGNEPKRLCHYDLYRIDQKNQEYIFQDAEEDDVLCIEWADRLPKPRFC DTINIYITMOTNMEREIIIEKR

CPn_0658 739180 738455 CT538 hypothetical protein KRWCMDISGAVKOKLLOFIGKOKKPELLATYLFYLEQALSLRPVVFVRDKIIFKTPEDAV R14EODKKIWRETEIQISSEKPOVVENTKRIYICPFTGKVFADNVYANPODAIYDWLSSC POMMEKOGGVRIKRFLVSEDPDVIKEYAVPPKEPIIKTVFASAITGKLFHSLPPLLEDFI SŠYLRPMTLEEVONOTKFQLESSFLSLLQDALVEDKIAAFIESLADDTAFHVYISQWVDT

CPm20659 739482 739838
t220-Thioredoxin
LOEDRADSNSIFREGKLMVKIISSENFDSFIASGLVLVDFFAEWCGPCRMLTPILENLAA
ELPHVTIGKINIDENSKPAETYEVSSIPTLILFKDGNEVARVVGLKDKEFLTNLINKHA

740327 739860

THE SOURCE OF THIS METHYLASE THE STATE OF THE STATE OF THE SOURCE OF THE STATE OF T

CÉRÈ0661 741139 740327
mip_FKBP-type peptidyl-prolyl cis-trans isomerase
H\$R\$ELKIKDRRKMNRRWNLVLATVALALSVASCDVRSKDKDKDGGSLVEYKDNKDTNDI
ELSDNOKLSRFFGHLLARQLRKSEDMFFDIAEVAGLQAELVCKSAPLTETEYEEKMAEV
OKEVFEKKSKENLSLAEKFLKENSKNAGVVEVQPSKLQYKIIKEGAGKAISGKPSALLHY KGSFINGOVFSSSEGNNEPILLPLGOTIPGFALGMQGMKEGETRVLYIHPDLAYGTAGO
PRINSLLIFEINLIQASADEVAAVPQEGNQGE

742938 741172

CPHE 0662 742938 741172

asps_Aspartyl trna synthetase
skGeymkyrthroreltsnhigenvolagwyrpyrnhggvyfidlrdrfgitqivcrede
QPELHQRLDavrsewlsvrgkvcprlagmenpnlatghievevasfevlskson/pfsi
ADDHINVNEELREVRYLDMRRGDIIEKLLCRHQVMLACRNFMDAQOFTEIVTPVLGKST
PEGAADYLVPSRIYPGKFYALPOSPOLFKOLLMVGGLDRYFQIATCFRDEDLRADRQPEF
AQIDIEMSFGDTQDLLPIIEQLVATLFATQGIEIPLPLAKMTYQEAKDSYGTDKPDLRFD
LKLKDCRDYAKRSSFSIFLDQLAHGGTIKGFCVPGGATMSRKQLDGYTEFVKBYGAMGLV
WIKNQEGKVASNIAKFMDEEVFHELFAYFDAKDQDILLLIAAPESVANQSLDHLRRLIAK
ERELYSDNQYNFVWITDFPLFSLEDGKIVAEHHPFTAPLEEDIPLLETDPLAVRSSSYDL
VLNGYEIASGSQRIHNPDLQSQIFTILKISPESIQEKFGFFIKALSFGTPBHLGIALGLD
RLVMVLTAAESIREVIAFPKTQKASDLMMNAPSEIMSSQLKELSIKVAF

CPn_0663 744220 742901
his3-Histidyl trna Synthetase
KSNHFERRHHYTVTLPKGVFDIFPYLADAKQLWRHTSLWHSVEKAIHTVCMLYGFCEIRT
PIFEKSEVFLHVGEESDVVKKEVYSFLORKGRSMTLRPEGTAAVVRSFLEHGASHRSDNK
FYYILPMFRYERQOAGRYRQHHOFGVEAIGVRHPLRDAEVLALLWDFYSRVGLOHMOIQL
NFLOGSETRFPYDKVLRAYLKESMGELSALSQORFSTNVLRILDSKEPEDQEIIRQAPPI
LDYV3DEDLKYFNEILDALRVLEIPYAINPRLVRGLDYYSDLVFSATTFTQEVSYALGG
GRYDGLISAFGGASLPAGGFGVGLERAIQTLLAQKRIEPQFPHKLRLIPMEPDADQFCLE
WOOHLRRIGIFPEVDWSHKKVKGALKAASTEQVSFVCLIGERENISQQLVIKNMSLRKEF

CET_0664 744775 744557 to cobust homolog present in Genebank/EMEL as of 11/7/98 LWCAHAMKKLIALIGIFLVPIKGNTNKEHDAHATVLKAARAKYNLFFVODVFPVHEVIEP

CTD_DBGS 74499R 746365

Obj. Hexospheriphare Transport

PMNWTKFFOFFKHIKEIEDDEVVKKKYKZWRIRIF/CMFIGYIFYYFTRKSFTFAMPTL

IABIAFDKAQIALIGETLYFIYGIGKEFVSOWSDGS/PRYFMAIGMITGLHIFFEMSS

IYIFALIAM BINGMFQ MCMPFCARELTHWYAKSEF/CTMMS/WETSHINGALIFILTOF

IIDYSCHIRGAMYVITILLIGMGLVLINRLRDTFQ/I/SLPPIEKYKRDPHHAHHEGKSASE

CTEELERELTTREILFTYVI/TNOWLMFLAAASFF/I/TVRMAVNDWSALFLIETKHYAAVK

ANFCVSLFEIGGLFGMLVA KISKGNEGEMNVI DGTLLFVIGFFLYGPOMMIGL-BAELSHKKAAGTASG WGWKGFFIALLACASIALLLFLFTWNAJEKNTRSKA KISKGNEGPMNVLEGLOLLFAILGMWESRSHNOWWV AFI.SHKKAAGTASGFTGWFAYFGATFAGYPLGKVTDV

746379

CPn_0666
dnaE-DNA POI III AIPha
GFFLTWIPLHCHSQYSVLDAMSSIKDFVAKGQEFGIPALALTDHGNLYGAVDFYKECTQK
GIOPITGCECYTAPGSFFDYKKEKRSRAAHHLILLCKNEGGYPNLCILTSLAFTEGFYYF
PRIDKDLER/YGEFILTELD VILLYCLIPAALKJEPAALLELAWFCDLFFCJYFTEVZLIK
CONSTRUCTOR OF CONTROL OF MILEGIAGPKEEMLEGETH. 1960 F PITAMERA BILATITYAMNEHETHANDAGA
EILLNVQSGETVRIAKQNTHIPAPKRKVYRSREYYFKSPAQMAELFKDIPEVISNTLEVA
KRCDFTFDFSKKKNYPIYVPESUKTLNSYTEEDRYQASAVFLKQLAEEALPKKYSSEVLAH
HAKKFPHOPIDIVKERMEMEMAIITPKOMCOYLLIVWDIIHWAKANGIPVGFORGSGAG
SVLLFLLGITEIEPIRFDLFFERFINPERLSYPDIDIDICMAGRERVINYAIERHOKDNV
AQIITFGTMKAKMAVKDVGFTLDMALSKVNHIAKHIPDLNTTLSKALETDPDLHQLYIND
AESAQVIDMALCLEGSIRNYGYHAAGVIICGOQLTMHIPICISKDSTMITTYYSMKPVES
VUMLKVDLLGLKTLTSINJAMSAIEKKTOGSLAMATLELDDATTFSLHQGKTMGIPQME
SKGMQELAKNLRPDLFEIIAMGALYRPGPMDMIPSFINRKHGKEIIEYDHPLMESILKE
TYGIMYYOGOYMQIAGALASYSLGEODVLRRAMGKKDFOQMEQEREKFCKRACNNSIDPE
LATVIFDKMEKFAAYGFNKSHAAYGLITYTTAYLKANYPKEWLAALLTCDSDDIEKIGK
LIREAQSMGIPILPPHINVSSNHFVATDEGIRFAMGAIKGIGRGLIESIVEERDHHGPYE
SIRDFIQRSDLKKVSKKSIESLIDAGCFDCFDSNRDLLLASVEPLYEAIAKDKKEAASGV
MTFFTLGAMDRKNEYPICLPKDIPTRSKKELLKKEKELLGIVLTEHPMDTVRDHLSRLSV
VLAGEFENLPHGSVRTVFIIDKVTTKISSKAQKKFAVLRVSDGIDSYELPIWPDMYEEQ
OELLEEDRLIYAJKVLDKRSDSLRISGRMMKDLSIVNENIIVECDQAFBRIKNQVQNSSF
TMSTSGKETKAKNNEYNENGHTQALAPVTLSLDLNELRHSHLCILKKIVQKHPGSRTLVL
VFTQDNERVASMSPDDAYFVCEDIEELRQELVTADLPVVITV

CPn_0667 751097 750177
No robust homolog present in Genebank/EMBL as of 11/7/98
NISLLCKIOKRYFMKKLILYFAAFVASLFCGVFLWDRVPCAQKIMRLAADHSSEVFSKSC
RFVRKISGFELQVFERHVSPEQALALFPEYRDGKSFVELAF IPHTLMHVRFSKEEPVKK HIISQEGETLWSLVNGEMYLHTGTWTCSKGFRECLLLHAGKODMRVIOTLATLGGTTSRE SLAQALALKNIRAERVIKECOKKKLIFASGNQIGTHFQOFQPIRGCTTTLNNNPVWLQKP RHAAVFPAQYSEDRVRHLVKMIFGDNFLIVRSSMVYVPVYKISLVSADNSVRVEYINAVT GKSFORL

CPT_0668 751176 752162
CT547 hypothetical protein
WRFWVVSPRLIMKFLLYVPLLLVLVSTGCDAKFVSFEPFSGKLSTGRFEPQHSAEEYFSQ
GOFFLKKGNFRKALLCFGIITHHFPRDILRNQAQYLIGVCYFTQDHPDLADKAFASYLQL
PDAEYSEELFQMKYAIAQRFAQGKRKRICRLEGFFKLANADEDALRIYDEILTAFPSKDL GAQALYSKAALLIVKNDLTEATKTLKKLTLOFPLHILSSEAFVRLSEIYLQQAKKEPHNL DYLHFAKLNEEAMKKQHPNHPLNEVVSANVGAMREHYARGLYATGRFYEKKKKAEAANIY YRTAITNY POTLLVAKCOKRLORISKHTS

752775 752140

CPH_0669 7.2140 7.2775
CT548 hypothetical protein
IEYLSILPKIEINMRLFSLGTIYLFFSLALSSCCGYSILNSPYHLSSLGKSLLQERIFIA
PIKEDPHCQLCSALTYELSKRSFAISGRSSCAGYTLKVELLNGIDKNIGFTYAPNKLGDK
TRAFFIVSNEGRLSLSAKVQLINNDTOEVLIOQCVARESVDFDFEPDLGTANAHEFALGQ
F#MHSEAIKSARRILSIRLAETIAQQVYYDLF

Pn_0670 752738 753196

rsbW-sigma regulatory factor-histidine kinase PRRLLNRYTMTFFEGETVFPAVLSELHSMLDLIKRAGKQSKCPQEKLLKLELACEELLVN IISYAYQGENSPOTIAISCISRRGDLEVVIKDHGPSFNPLAVSINIQEDLPLEQRKLGGL GIFLAKSSVDEFLYAREDHCNIVHLKMLNGQHS

CPn_0671 753660 753205 CT550 hypothetical protein RITINQRKYTMSLDFFEEFYHOSILNTGTSFPEGYLNIAEILSYPHCTDANTDFLCSQSD MDFIIAESKDKLTLFNADFAIWLYPELVQGQAVTRGYIAVSQGEGNYEPEMAFEASGQYN QSSLILEALQLYLKDIKDTENALRSFRFNNDH

753723 755048

CPM_0672 753723 755048
dacf(pbp5)-D-Ala-D-Ala Caroxypeptidase
TIKSPHMKRPFFTYLCIIFYGSCASLSLHAGLSFPEVRGATAAVVHADSGKVFYDKDIDA
VIYPASMTKIATALFILKHYPTVLDTLIKVKQDAIASITPQAKKQSGYRSPPHMLETDGS
TIQLHLREELLGWDLFHALLUCSANDAAVVLAMACCGSVEKFMDKLNFFLKEEIGCTHTH
FNNPHGLHHPNHYTTTROLISIMRCALKEPPFRGVISTTSYKIGATNLHGERILSPTNKL
LLPGSTYHYPPALGGKTGTTKTAGKNLIMAAEKNNRLLVTIATGYSGPVSDLVQDVIALC
ETVPNEFLLRKELVPPSDCLQLEIANLCKLSCPLPEGLYYDFYASEDREPLSVSFIAHAA
AFPIEQGDLLGHWVFYDDEGKKISSQPFYAPCRFERTIKPWKLYMKRVFTSYRTYMSITM
LIMYEPIEUBVYNIN UNDSYL LLMYFRIRKHRKYKNLKHYSKI

CPn_0673 755242 755463
CT552 hypothetical protein
GKSTEGKAYHCFLKQVSIALNREEVWDNPHHLMFILMQFQQFSGEQDRFGSFLEATIRDR
VSFLVLQEKIATLK

756689 755577

CPn_0674 75689 755577
fmu-RNA Methyltransterase
RGILYVTMVPFROHHAYQLLKQLHTSAISEADRVSYYFKQNRSLGSKDROWIQNIIFNIL
RHRRLLETLILDSGEQVTPEALVAKVNEGVLENLDSYSAIPWPVRYSISDDLAHFSKHHPUQSY
GEGOAEEIAKIMLTEAPTITRVTHOKIGVKELQEKLEYPPSPGELPEALHFSKHHPUQSY
EAFRRGFFEIQDENSQRISQJICHTDKDIVLDFCAGAGGKSLIFAQKAKHVVINDSRKAI
LQTAKHRLLRAGARNFSLAPQLPLDSFCVVIVDAPCSGTGVFRRHPEHKWQFSKKLLLNY
VRVUKSILKQASAYVGPRGRLVYITCSLLKEENEAHVAYMHSLWKEVHRKTLPLQVGKG
DAFFTSHFQKI

757931 756768

CPR_0675 75/931 756768
CT696 hypothetical profess
VPLSMILDFORSIGYYLRVLELAIRDFFRILAYDRKRULLDAWPYNDFUPTNYDTSVSTI
RQVIHELFSWSAILSYSISSRILAIIFEDFUHERKFGROWD/RUPFUPTNYHIKKAIVORLCM
RKGULLFESRRPVKIVQAANKYEGKORONESSWEDFFHEEYTVSEWAPFLAIGEVQRELAA
DAGUMIIEALTTULEGHTAYUFLEHLINGFIGEKAQFEKTLSEKGYVILRELIQUFSL
SAEDFOTIIMSIISDGUSEVLAHELIFERDFIFEKRFFWILWGETALASPEDSKIALGFU
AEVLRKVIVEKKLHVGKSPNTTPEEVGHTYGIIYGIBUJNPALWDKMITMELMRWILDYDRDIG
IALRKAAEYYNPHPSFWRGFLPLWOPPF

CPn_0676 759220 758051
homologous to CT695
SMCTPISGNDGDRNTISDPLEESAAEEGDSDLEDRVSESATOVIETIADTGIPEATPSEG
SMCTPISGNDGDRNTISDPLEESAAEEGDSDLEDRVSESATOVIETIADTGIPEATPSEG
TNSDLNDGLVDRVEYEARGSLLTTMLAR IRKAVSQIWMHVKTKRHPKEQGVRSLGDIPCD
LLKATRLPKETAEPPYFYALETALASCRSFFFHVFLRLFTLLRRQHPEAPLDLGGTDPIS
PEAAVAFALILRSCCKWATDAVQECLPLEVIEEAGWYNAFSLEATTTVEEVSKRLSELL
YSDKRIGGLANVRGITKIITSFYLGAGQCVSVVONLKTYDLGRNYTQVLACASQIDEFAD
KGENEALVMKDILYIJVRORGKEIJDFLMMWSEEHASEVNYDVVLAILEVNLPILEEDYR
UHHAAVGEKLRIVETQEFLIGERLIGHEERD

CPn_0677 760410 759256

No robust homolog present in Genebank/EMBL as of 11/7/98
RIAMGINPSONRSPDDVWVRGAQGDSSSTQGTGATNSNLGAHNVTTSTSQPOVASKAKQL
WQTVREFFLCKKSPDSSQCASGPAMQSPSGPTIRFTRAPPPPTTGGANAKRPATHGKGR
APOPPTAGSSSGSEQPTAMSSEVAKLVSELKDAVHSHAESQKVLKKVSQELQTKWTDWEN
NRGPDYLLHGYRVIARALQQTTTEOSMLIEGTSSTGPVPQAVTVAKDAVTQTVRGAIKNL
ENPKPGNDPDGVLMQVVISLGIEGPPLDPGESIONFLETRVSDFGGDDSDIDYTSDIARL
GSALDRVRENHPNEMPRIWIALARELGAVHSHATSVRIANAGKNHTRDVVRMANESSRL
LOCKKKUSVGAWANTMTVLIGDI.FE LOCMKVLSVGAWANTMTVLIGDLFE

CPn_0678 761329 760682
No robust homolog present in Genebank/EMBL as of 11/7/98
KIIMSVNPSGNSKNDLWITGAHDQHEDVKESGVTSANLGSHRVTASGGRQGLLARIKEAV
TGFFSKNSFFRSGAPRSGQPSAPSADTVRSPLPGGDARATEGAGRNLIKKGYQFGMKVT
IPOVPGGGAQRSSGSTTLKFTRFAPPPPKTGGTNAKRPATHGKGPAPQFPKTGGTNAKRA ATHGKGPAPQPPKGILKQPGQSGTSGKKRVSWSDED

CPn_0679 762936 761725

pgk-Phosphoglycerate Kinase
gympkltvodlspeekkvlukvdprnyphodkillddirirsamptinyllkkhaavilms
HLGRPKGGFOEFSLGPVVDHLGSVLGHHVPLAPDCVGEVARQAVAQLSPGRVLLLENL
RFHIGEEHPEKDPTFAAELSSYGDFYVNDAFGTSHRKHASVYVVPQAFFGRAAGLIMEK
ELFFLGRHLLTSPKRPFTAILGGAKISSKIGVIFALLNQVDYLLLAGGMGFTFLQALGKS
LGNSLVEKSALDLARNUKIAKSRNYTTVLPSDVKAAENLGSKEYSVISIDQGIPPHLQG
FDIGPRTTEEFIRINQSATVFWNGPVGVYEVPPPDSGSIAIANALGMHPSAVTVVGGD
AADVAL AGCGTXVENDTGCGALSLEFFLGGERIEDTEFULSGEKS AAAVVALAGCSTKVSHVSTGGGASLEFLEQGFLPGTEVLSPSKS

764254 762971 CPT_0680 764254 76291
ygo4-Phosphate Permease
YSMLPLIIFVLLCGFYTSMNIGANDVANAVGPSVGSGVLTLRQAVVIAAIFEFFGALLLG
DRYAGTIESSIVSVTNPMIASGDYNYGMTAALLATGVWLQLASFFGWPVSTTHSIVGAVI
GFEEVLCKGTIIYWNSVGIILISWILSPFMGGCVAYLIFSFIRRHIFYKNDPVLAMVRVA
PFEAALVIMTLGTVMISGGVILKVSSTPMAVSGVLVCGLLSVIITFYVYHTKHCSYISDT
PKRGSLTYPLKERGGNYGRKYLVVERIFAYLQIIVACFMAFAHGSNDVANAIAPVAGVLR
QAYPASYTSYTLIRLMAFGGIGLVIGLAIWGMRVIETVGCKITELTPSRGFSVGMGSALT
LAZESILGLPISTTHVVVGAVLGIGLARGIRAINLNIIKDIVLSWFITLPAGALLSILFF FALRALFH

CH10681 765001 764258
CT691 hypothetical protein
nG##SHKSPTRSFRQVIIAKKAILMOTLARLFGQSPFAPLQAHLEMVVSCVEYMLPIFTA
LRDGRYEELLEMAKLVSDKEYQADCIKNDMRNHLPAGLFMPISRAGILEIISIQDSIADT
AEBMAILLHIRRLNFYPSNETLFFFRILEKNLEAFELTMTLLHEFNQLLESSFGGRKADKA
RLIAGRVAKSEHESDVLQRELMQIFFSDDFIIPEKEFYLWLQVIRRTAGISDSSEKLAHR INMTLEEK

CPh_0682 764912 765955

dpb_ABC ATPASE Dipeptide Transport
TSKGLHKNSLFRNNNLPKRSCKRLMASNPILOIEDLSITLAKORQOYPIVOSLSFTINEG
OTLAIIGESGSGKSVSAHAILRLLPCPPFSVSGQVNFQGHNLLTASRSIQKKIIGTEISM
IFONPQASLMPVFTIEQOFREIHTHLALTAEVAKEKMLYALEETGFHDPRLCLNLYPHQ
LSGGMLQRICIAMALLCSPKLLIADEPTTALDVSVQVOILGLKHTLOKKTGKSLLIITM
MGVVAETADDVLVLYAGRMVECAPAVQMFHNPSHPYTRDLLASRPSLQPQQLGSFNPFG
QPPHYTAFPSGCRYHPRCSKILNRCSAEAPEIYPVREGHKVRCWLYDD 764912 765955 CPh 0682

765936 766919 CPHE_0683 765936 766919
dppR_ABC ATPase Dipeptide Transport
GVGCMTINFPQPLIQATSLTKHYYKRSFWFQGKTIASRPVDDVSFSLYSRRAVGLIGESG
SGKSTLALALAGLIPLTSGFLIFTNOTPIKLHSKHGRHQLRSQVRLVFONPQASINPRKTI
LDSLGHSLLYHKLVPKEKVLATVREYLELVGLSEEYFYRYPHQLSGGQQQRYJIRRALIG
VPQLIICDEIVSALDLSIQAQILNMLAELQKKLSLTYLFISHDLAVVRSFCTEVFIMYKG
QIVEKGNTKKIFSDPQHPYTRMLLNAQLPETPDQRQSKPIFQEYHKDSEESSTGCYFYN RCPOKOEACKSEIIPNOGDAHHTYRCIH

768056 767181 CPD_9684 /68056 /67/181
spoJ/parB-Chromosome Partitioning Protein
EKSGDIVTEEISKDTIIEVAIDDIRVSPFQPRRVFSNEELQELIASIKAVGLIHPPVVRE
ICTGDRVLYYELIAGERRWRANQLAGATTIPVILKHVIADOTAABATLIENIQRVNLNPI
EMAEAFKRLIHVFGCITODKVAYKVGKKRSTVANYLRLIALSKTIQE\$LLQQITLGHAXV
ILTLEDPILREKLNEIIIQEHLAVREAELIAKQLISEEGSSIELKFTPLDMAESSKQHEE
LQQRLSDLCGYKVQIKTRGSKATVSFHLQNTQDLQKLEAWLSSHGTLSESLS

CPN_0685 768026 768217
No robust homolog present in Genebank/EMBL as of 11/7/98
FPOSQYLLIFPNRILDLQAFEILDVQGMLTDQRKHIQMLHKHMSIEIFLSNMVVEVKLFF

768373 768176 CPn 0686 ten_0000 No robust homolog present in Genebank/EMBL as of 11/7/98 AKDSMMPGGRLFRVYQELFFFSSVYVCEQRRPRKLYPSLOHLNFPIEKPRFLLKGFKKEL

768501 769214 CPE_9687 768501 769214
CT482 hypothetical protein
RK HKRILMHAYRESTENGESEMEKLUHNIWEKEYSESSATATCIVLASELSEK IVSNTYK
HSOAKENSILLUTRAAEVAVSOGELPSKSALSSLEGAYHLGESMKPYAGELASCEYTHN
EPLIKAYYASLAYRISQALQLEHFIQKLIKETSEAQADQLYDVALSKSYQLLQTANSSPE
YRTLSELTLIKVIELKELLHQDVSQDFAALKSSELFHOFERMYSDGEWTLSKREGKKG

76.2376 CPn 0688 CT481 hypothetical protein SUMLIVLAFROVFFSHSRS NYLBÍLIKONFAITUPKERTIKUHUIMLTFEFASFD FYTNIFPFLEEQKIPAVVIV 19SNAQOLIFSHRIKUHETLAFDDEIFSNYMPFCC ONELIEMAKSPYIQLASSGFAIRNLMNYPPYLTTEILLSRHHIETITJAKPLAFLFFFGK SDPTSRKLAADHYPYSFLIGNTINRKIKTHNIYRLDIKPMQYVCPSLFQSSRYLKNWIKE KSKQLYLKKQLPKR

770197 771407 CMN_U059

Vfh0-Nifs-related Aminor cansferase

VMN_KEPPPIFAMMAKEDEPTIL AMATORES

SHITEAYAAVREPPPHAMMAADUSELUSTE STO .vipavameye nyaenea iyusi HARLAL VNEUMIERO HHANVLSWEIACRRGSLVKKJRVHDSGL IDLDDLEKLLNEGAGFVGIPHVSNVTGCVOP LQOVAELVHRYDAYLAVDGAGĞAPHLPIDVQLWDUDFVYFSSHKIYGETGIGVLYGKKDL LDQLPPVEGGGDMVAIYDHQVEFYLPAPMKFEACTPNIAGVLGLGAALDYLDGLSAKFIY DKEIALTTYLHKELLEIPGVEILGESIEEPRGALIGMTIDGAHPLDLGFLLDLRGIAVRT GHQCAQPAMERWNVGHVLRVSLGIYNDEDDIDQFILVLQDSLDKIRR

771436 h72704 CPn_0690 /172704 771436
ABC Transporter Membrane Protein
LSVLRGDKVLVSIETF\$SIASGSPVOKAAEACYTQYSKQPSSKEVLSSFSWIQELSLFPD
RYNLATGASELIKOHVLHNNHSLAFECILINGKYEPSLSQLPBGVIVCGIDEARGSLSSF
WGGFDVNKHPLAFLAVCSEDGRVIYIFDEEMQTSDPIFVRHISFPTVSDHDVIFSPRIV
VILQRASAQIQISADVDLEBWGSSKTIVNGVTELFVGEGADLTVTMVPGYSEEDTLSWS
TIATVEKDAICRMONLLESCOGFGWFDNTSYIVGKKGHAESLVLVQSPRKTWANNLMS
DAEETVSRQNIKSILVSCHFLFECTISISSCODLSDANGKHDTLLLSSEARVSTFPRLEI
ETDEVKASHGATVGPLDPQQIFYMRSRGMTEAEAQEKLIHGFLKQGLVSDTFLGSSFQLN

773467 772685 CPn 0691 CPT_0691 //346/ //2005 CT691 hypothetical protein RGLGSMLK/KHLHASCNDVKILDDFNLNIOPGTMHVIMGPNGAGKSTLAKILAGDESVLV SSGEIALOGONLLSMLPEERSRAGLFVGFQMPPEIPGVNNKMFLRDAYNARRRANQEGDI SIDEFNTLISTVLETYENNATTDLFLDRNVNEGFSGGERKRNEICQMLVLEPEMVLLDEP DSGLDVPALRLICRVLEKKRELHPTSSLCIVTHNPKLGNLIRPDVVHLLLDGRVALSGDV SLMHELEAKSYQEVTKRVAWR

ø692 774945 773461 ABC Transporter
IQBFCATGLKVMGESVKVFLEEREDYPYGFVTPIESQGLTRGLSEETIEEIAALRNEPQF
IQBFCATGLKVMGESVKVFLEEREDYPYGFVTPIESQGLTRGLSEETIEEIAALRNEPQF
IJBFRLQAYRYMKQLHEPAWARLHYGPIAYDDIVYFSSPKQKKPLGRLEDADPEILDTFK
KLIGIPLDEOKRLLAVENVAVDLVFDSVSIGTTFKEALEKAGVIFCSLGEAIQEHPRLVKK
YLGSVVSHRDNFFAALNAAVFSDGSFVVPKGVKCPMDISTYFRINNKEAGQFERTLIV
EDGGYASYLEGCTAPAYSSNQLHAAVVELVAHEHAVIRYSTVQNWYAGDKKTCKGGIYNF
VTKRGLCAGYRSKISWSQVEVGAAITMKYPSCILKGDESVGEFYSVALTSGKMQADTGT
KHVVGRNTSTVISKGISSDESKNFTFSLVSLGKKAEHSSNTTQCDSMLIGKASGAYTDP
KIVVENSTSSIEHEATTSKLREDQLLYLRSRGLSPEEAVSLVIHGFCREIIEQLPLEFAQ ABC Transporter EASKLLLIKLENSVG

775240 _0693 776292 CRT_0693 776292 775240

TPR Repeats (0-Linked GlcNac Transferase homolog)
LIGSTNHVLGEISMEEAAKHLAKEFLCSGINLFLSGEYEQAEKRLKETLELDSTAALAYCY
LIGITALETGRVSEALNWCSKGLASEPGDSYLRYCYGVALDRGNQYEAAIEQYSAYVALHP
DDVECWFSLGSVYHRLKRLQEALDCFDKILALDPWNPQSLYNKAVILSEMDDEAESIRLL
EVAVAKNPLYWKAWVKLGFLLSRSKRWDKATEAYERVVQLRPDLSDGHYNLGLCYLTLDK
TRLALKAFQEALFLNAEDADAHFYVGLAHLDLKQMREAYEAFNSALSINLEHERAHYLLG
YLHHMQGETDKATKELLFLQKKDSTFAPLLQKTVVSDPSSMQFERRLDTIS

CPn_0694 779635 776330

pbp2-PBP2-transglycolase/transpeptidase
FSDESEAHNIHSMKRPKKFPIYLSIAOKTNRLLSGIVTAFAVIALRLWYLAVVEHEQKLE
EAYKPQIRVLPGYVERATICDRFCKTLAVNQLGYDVSVAYGAIRDLPTRAWRVDEHGHKQ
LIFVRKHYIMCLSELLSOELHLDREAIEDAIHAKASVLGSVPYLVAANVSERTYLKLKML
SKDWPGLHVEAVVRHYPQESVASDILGYVGPISLQEYKRVTQELSQLRECVTAYEEGEF
FKLPEGLASIDQVRALLESVESNAYSLNALVGKMGVEACWOSKLRGKIGKKPILVDRRGN
FIQEMEGAVPEAPGTKLQLTLSAELQAYADALLLEYEKTETTFRSAKSLKKREKLPPLFFW
HKGGAIIALDENNGEILAMASSPRYNNDFVNAKVAEDSKAVRSSIYRMLENKEHIAEIY
DRKVPLIRERRNPLTGLCYEEILPLTFDCFLDFLFPENSVIKLQLKRNSFVGQAIEVQNL
VTRLLSLFPYEEGTCPCSAIFDAVFPNEEGHILIQEVISLQEDKWIMECLNQHKADIEEL
KEALDQVPNEEGTCPCSAIFDAVFPNEEGHILIQEVISLQEDKWIMECLNQHKADIEEL
KEALDQVPNEEDTAVPKSWRKSEFLQYLAAKRQEBALRKORYPTPYVDVLEEEKTRQYKM
FCQEHLDTFLAYLFSKTPYKEGLEPYYDILDLWINELDNGARALSWHEHYLFLKERVSH
FCQEHLDTFLAYLFSKTPYKEGLEPYYDILDLWINELDNGARALSWHEHYLFLKERVSH
FCQEHLDTFSTFREFNELQRELLGKYPISIVTNKROTEQDLAASFYPVYGYGYLRPHAYG
OAATLGSIFKLVSAYSVLSORILWGHNEEPANPLVIIDKNSFGYRSSKPHVGFFKDGTPI
PTFFRGGSLPONDFMGRGFIDLVSALENSSNPYFSLLVCEGLGDPEDLADASSLFGFGEK
TGLGLPGEYAGRVPHDLAYNRSGLYATAIGOHTLVVTPLQTAVMLASLVNGGVYVYPKLL
LGEWEGEHVSYLSSKKKRTIFMPDAVVEVLKTGMRNVIWGQYGTARAIQSOFPPQLLSRI
IGKTSTAESIHRVGLDBEFGTMKMKDIWFAAVGFSDQDLSLPTIVVIVILRLGEFGRDAA
PMAVKMIDMWEKIQQRESFLRG 776330 779635 PMAVKMIDMWEK IQQRESFLRG

780201 781382 homologous to CT695 homologous to CT695
SLEVSMKKLLKSALLSAAFAGSVGSLQALPVGNPSDPSLLIDGTIWEGAAGDPCDPCATW
CDAISLRAGFYGDYVPDRILKVDAPKTFSMGAKPTGGAAANYTTAVDRPNPAYNKHLHDA
EWFTNAGFIALNIWDRFDVFCTLGASNGYIRGNSTAFNLVGLFGVKGTTVNANELPNVSL
SNGVVELYTDTSFSWSVGARGALWECGCATLGAEFGYAGSPKVEELWICDNVSCPSVNK
PKGYKGVAFPLFTDAGVATATGTKSATINYHEWOVGASLSYRLNSLVPYIGVGWSRAFTFD
ADNIRIAQPKLPTAVLNLTAWNPSLLGNATALSTTDSFSDFMQIVSCQINKFKSRKACGV
TVGATLVDADKWSLTAEARLINERAAHVSGQFRF

781703 782599 CPG_0696 781703 782599
CT696 hypotherical protein
nddfyvrmplltysnfelevosledobsklifikdlmdagahfydigtrwnpkmklyifee
knglyiinlaktlogleiaaliaiirwiidbnkfylevofykkoakuvirkaaleadeffiae
kwlddkthnmttrknifkriidbiekolashoayitykkeaaleakhiloklikhiledirknik kapallovoodsykkiavaeakkiviiivlaloothkodyffidhulisinddalksirklik A LA ENT LE EMBET MULE LA MANCHEA LO DES DE ENBEED PLY PRED DE EN

782587 783447 ETARNITOWALTON LEGICOL LC.
MANIKINODESWELFRIFACTOL LC. SEQUESTIVEERAVING TVOERH RESERVAY FEKATISTVOEYSHOM SKIVALIMILISISS

TADSLAKDIAMHVVAAQPQFLSKESVPAEAIAKEKEV FFQEACLLEOPFIKNADLSIQSLIDDFSKTSGSSVAI

CKPOEV LEK IVTGKLNT

783443 784201 t*Pn 0648

CPD_0949

DYH-UMP KINASE

EPNKNMAKQTRRVLFKIGGEALGKDSSNRIDEMRLSRLVSELRAVRNNDIEIALVIGGGN
LURGLAEQKELQINRVSADQMGMLATLINGMAVADALKAEDIPCLLTSTLSCPQLADLYT
POKSIEALDOGKILCTTYTAGSPYLTTDTGAALPACELNVDVLIKATMHVDGVVDKDPRL
#FDALEYDFYGYEBFLGRUGFMAGAGAEGEMDGHIFIRVFGFUJHDGKFALFDFTITTL SORDAMHAGGGREE

784179 784721 CPn_0699

CPT_0699 /M4177 / APA/AL
rtf-Ribosome Releasing Factor
TMSVLQDTEKKMAAALDFFHKEVKSFRTCKAHPALVETV-VVDVYGTTMRLSDIASISVAD
LRQLVISPYDGNNASAIAKGIIAANLTQPEVEGSIIRIKVPEPTADVRQEMIKQLRRKC
ERAKINVRNIRREANDKLKKDSALTEDVVKGNEKKIQELTDKFCKQLDELTKQKEAEIAS

CPn_0700 785094 785609 CT676 hypothetical protein LMYHSPTHQCYHCQQPATICYTEIDKDKVIRSYVCATCPCPSHYYNNEHLSLSKGVGVLT LECGNCKTVWHSKQDDEQLLGCHQCYTNFKNQITSKLKSERVVSSSFTMEKGQGSLHIGR APGEASNTNPLLKLIALNEALQDTLEREDYEQAAVIRDQINHLKTKNPDDPS

CPn_0701 785584 786672
karG-Arginine Kinase
KPKIQNTLPNDLLETLVKRKESPQANKVWPVTTFSLARNLSVSKFLPCLSKEQKLEILOF
ITSHFNNIEGFGEFIVLPLKDTPLMQKEFLLEHFLLPYDLVGNPEGEALVVSRSGDFLAA
INFODHLVLHGIDFQGNVEKTLDQLVQLDSYLHSKLSFAFSSEFGFLTTNFKNCGTGLKS
QCFLHIPALLYSKEFTNLIDEEVEIITSSLLIGVTGFFGNIVVLSNRCSLGLTEELLLSS
LRITASKLSVAEVAAKKRLSEENSGDLKNLILRSLGLLTHSCQLELKETLDALSWIQLGI
DLGLIKVTENHPLWNPLFWQIRRAHLALQKQAEDSRDLQKDTISHLRASVLKELTKGLSP

CPn_0702 789700 786929

yscC/gspD-yop C/Gen Secretion Protein D

LKKNPVXTVILNIGRKILQGIKKKKKKIGILSGLFFLDLVLLGVSSQRPTETSANVKHNL
RDEKLAACPKNSAASLSAKKSHTKKTTPGSIFSKT/SKTPATQDKTFKTSGSAFPAKPT

TLKELEERKKPRPERRTTADVKRSPFLPTOEVEEPVPAASKEQLDSIQWEEKQNYARR
AVNAINLSIKKQLEEQTSTVTEKDVQPKTQATPHASKKNVASPSTSMFGIEKAATTVAVP

QDKSEEEKVKERLTKRELTCEDLKONGTVTVFEDISIELLGFVSKISGTNFVFSDSLQ

FINTIVSHDPTSVDDLSTILLQVLKNHPLLKVVEQGNNVLIYRNPHLSKLSTVVTDSSLKE

TCEAVVVTRVFRLYSVSPSAAVNIIQPLLSHDAIVSASEATRHVIISDIAGNVDKVSDLL

AÄLDCGTSTSVMTAMTEDDFATALALGGTGTTTSPKSLRFFMYKLKYQNGEVIANA

LQPEGYNLJVTTAMDEDFINTINSIQWLEVNNSIVIIGNGCNVDRVIGLLNGLDLPPKOV

YIEVLILDTSLEKSWDFGVQWALGDEQSKVAYASGLLNNTGIATPTKATVPPGTPNPGS YIEVLILDTSLEKSWDFGVQWVALGDEQSKVAYASGLINNTGIATPTKATVPFGTPNGS
IPEFTPGQLTGFSDMINSSSAFGLGIIGKULSHKGKSFLTLGGLLSALDQDGDTVIVLNF
RIMADDTQQASFPVGCTVPYGTTNTIIQFETGTVTONIDVEDIGVALVVTSTVAPNNVTLI
QTEGTISELHSASGSLTPVTDKTYAATRLQIPDGCFLVMSGHIRDKTTKVVSGVPLLNSI
PLEMBGLFSRTIDGRQKRNIMMTIKPKVISSFEEGTRVTNKEGYRYNWEADEGSMQVAPRH
APECCGPPSLQAESDFKIIEIEAQ

789685 791205

CPE_0703 791205 789685
pkis-s/T Protein Kinase
RKIGFMCCRGIELPEPQVIGGYHVKKILSKKLRSRVVHGLHPETRHSTVIKVFSPSSF
TSRSVYNFLKEAGSLHGITHPNIVKFHRYGKWODCLYIAMEYIEGISLREVILAGFÉSLP
QAIDIIFDIAQALEHLHSRNILHKDIKPENILITPQGKIKLIDFGLADWDTEIQRAHPSV
GTFYYMSPEQRGESHSPASDIYALGLLAYELILGHLSLGRVFLSLVPERISKILAKAL
QESENNRYSSTREFIODIHHYRNSGDMCEDLRIKDHTVALYEQLGTORFWLAPFILFPD
FISGVLYHQGYPLYPHAYDTLLEGDVFNLWLGYSPISNATIALSVVKSLVCOGDLQRPLL
DRYGEINECLIRKKIPIDEMGISILCLEISKENKELSWIACGKTVFWIKRQGRVVQDFES
FSPGLGKITSLQIRETKVAWEIGDEAVVCTLELEESVASLKTLSLAELQDRQKAIFCPI
ESHGGIQSRQHGSNSPSTLISLKRIR

CHA 0704 792330 791209

flin- Flagellar Motor Switch Domain/YscQ family
RYEMAVANDSSASMLKSRNNFLSSLGKTEEQVAAPEFFRELCOHKIJEKFFLEDVQVSIK
FRGSITAVEATKEEGVHLLIOPMVOPMEVENLLFLTSEEDLGELWAVFDDASLASVFY
EKDKLLGFHYYFVAEACKLFEELQWYPSLSAKVGGDAIFTATSLGSFQVVDISLRLDGK
NVRCRLLLPEDTFOSCQKFFSGLHDESDLHNIDQTOQISLSVEVEYSQLTQEEWHQVVPG
SFIMLDSCLYDPETEESGALLTVQKHQFFGGRFLTPSSCEFKITSYPNLTHEDPPLPEND
QASAAPLPGYSRLVVEVARYSLAVSEFIKLNLGSILSLGNHPAYGVDIILDGAKVGRGEI
LAIGDUG GIBVLEV IALGDVLG IRVLEV

CPn_0705 793176 792334
CT671 hypothetical protein
FMELKKTAESLYSAKTONHTVYQNSPEPRDSRDVKVFSVEGKQTRQEKTTSSKGNTRTES RKFADEERRVDDEIAEVGSKEEEQESQEFCLAENAFAOMSLIDIAAAGSAEAVVEVAPIA VSSIDTOWIENIILSTVESMVISEINGEQLVELVLDASSSVPEAFVGANLTLVQSGQDLS VKFSSFVDATOMAEAADLVTNNPSQLSSLVSALKGHÓLTLKEFSVÓNLLVQLPKIEEVQT PLHMIASTIRHREEKDQRDQNQKQKQDDKEQDSYKLEEARL

CPn_0706 793689 793180/ CT670 hypothetical protein YAVAKYPLEPVLAIKKDRVDRAEKVVKEKRRLLEIEQEKLREKEAERDKVKNHYMOKIQO

TAVAT FEBEVELARMEN OF THE TAVATOR OF

3704

CPn_0707 795035 793704
yoon-7op N (Fladellat-Type #Phase)
ymmolittopditimgoladunutityvortitevyorlikavypnyrygevolvkrngmepl
ytevyoptoefaplisploedsyspsseviptglplhiragniglisrvlinglepidvet
ygfigifixytepi-ferappoplirakkrolestryroloddevansorigifagagoks
ollicibil hanaeradynyn ylalicerskevybetigodijebimkei yvyyetodyssolen
aayyota (Aeyfrodoktyvlimmosytraralreyglandepparakytpoypstlpri
leeryaadkytpiapytyvlinavjöynepyadeyks) (Lochtylehalagayhypaidyla
eleklita (ypergrei lokapeylakykanemi ir igeyprochreidaladidhidking
elekgitulektnypeaakylikavykanemi ir igeyprochreidaladidhidking FLKODTHEKTNYEEAAQQLRA**Y**LFR

CPn_0708

795034

CT668 hypothetical Load April Aparthyreform idevects and Aparthyreform idev CT668 hypothetical AFKTVKRFFCFMIDPVECF

795203 CPn_0709 796203 CT667 hypothetical prote

ERGODERGAR MERONGARELAN NIFYCHOLOGALES MARGODINME COURGODERGARECUEC GERENO VEREIMAL LEVEL

QVIANCKIESTRALAQSVLLWHOTLVAKSAUPLU

738482 796210

CT666 hypothetical protein RSRGERSMATNRSCTAFDFNKHLDGVCTYVKGVQQYLTELETSTQGTVDLGTMFNLQFRM QILSQYMESVSNILTAVNYTEMITMARAVKGS

796791

CPn_0711 / 796791 796486
CT665 hypothetical protein
TTINNQVLGFINYLYLGRYSMFNMENTAKEEKNSQPLLDLEQDMQDHDRAQELKASVQDK VHKLHALLREGSDIJESFGQQQSLLAGYVALQKVLGRINRKMI

CPn_0712 / 799315 796781

FHA domain:/homology to adenylate cyclase)
MAVRLIVDEGPLSGVIFVLEDGISWSIGRDSSANDIPIEDPKLGASQAIINKTDGSYYIT MAVRLIVDEG/LSGVIFVLEDGISWSIGRDSSANDIPIEDPKLGASQAIINKTDGSYYIT
NLDDTIPIVWSVAIQETTOLKNEDTILLGSNQYSFLESDEFDPQDLVYDFDIPEENFSND
SGDLSDSNGGKDLEPPGTSETMISPKPKEKLIKTUGGSSDPITSGDQELADAFLASAKAE
KNQPRAKVAKKGLKESSNESLNPKEQNAKDSPKGEERTNKPQNAIMEDNGASPRQDPQPK
SAEPSLKYTARDETPLKENKEVEEKANKKATPDSPEKKOPEEGSKKEGSKIEATPLDSQ
KESEDKAEEAFVQEEEENITEDNKEDSDSAADANDDTASDHTAEDNKETPKKVENEKSA
VLSPFWQDLFRFDQTIFPAEIDDIAKKNISVDLTQPSRFLLKVLAGANIGAEFHLDSGK
TYILOTDPTTCDIVFNDLSVSHQHAKITVGNDGGILIEDLDSKNGVIVEGRKIDKTSTLS
SNQWALGTTLFLLIDHHAPADTIVASLSPDDYSLFGRQQDAEALERQEAQEEEEKQKRA
TLBAGSFILTLFVGGLAILFGIGTASLFHTKEVVPLENIDYQEDLAQVINQFFTVRYTFN
KTMSQLFLIGHVKNSTDKSELLYKVDALSFVKSVDDNVIDDEAVQEDMILLSKRPEFKG
ISMHSPEPGKFIITGYVKTEQAACLVDYLNIHFNYLSLLENKVVVETQMLASHLAGHLLQ
GGFANIHVAFVNGEVILIGYVNNDDAEKFRAVVQELSGIPGVRLVKNFAVLLPAEEGIID
KNLRYPNFRYTGYSRYGEISINVVVNGRILTRGDVIOMTVTSIQPNAIFLEKEGLKYK LNLRYPNRYRVTGYSRYGEISINVVVNGRILTRGDVIDGMTVTSIQPNAIFLEKEGLKYK

799332 799817 CPn 0713

CMT_U/13 // 1931/ CT663 hypothetical protein LDLKERKAGFRNEIVSIPQGTKTTIAALENTSMLEKLIKNFATYMGITSTLELDADGAYV LPISEVVKVRAQQNADNEIVLSASIGALPPSADTAKLYLQMMIGNLFGRETGGSALGLDS EGMVVMYRRSGDTTYDDFVRHVESFMNFSETWLSDLGLGKQ

801125 800091

CPT_0714 801125 800091
hemA-Glutamyl tRNA Reductase
NYRIVLMVLGVVGISYREAALKERERAIQYLQSFEKNLFLAQRFLGKGGAFIPLLTCHRA
ELYYYSESPEIAQALLSELTSOGIRPYRHRGLSCFTHLFQVTSGIDSLIFGETEIGGQV
KRAYLKGSKERELPFDLHFLFQKALKEGKEYRSRIGFPDHQVTIESVVGELLLSYDKSIY
TNFLFVGYSDINRKVAAYLYOHGYHRIFFGSRQOVTAPYRTLSRETLSFRQPYDVIFFGS
SESASOFSDLSCESLASIPKRIVFDFNVPRTFLWKETFTGFVYLDIDFISECVQKRLQCT KEGVNKAKLLLTCAAKKQWEIYEKKSSHITQRQISSPRIPSVLSY

CPn_0715 801636 803462

GYTB-DNA GYTASE SUBURIT B

KFNKISHMAAYTEASILSLASLDHIRLRAGMYIGRLGMGSQKEDGIYTLFKEVVDNGIDE

FIMGHGKSLKISASDKQISIQDQGRGIPLGKLIDCVSKINTGAKYTQDVFHFSVGLMGVG

LKAVNALSEIFSVRSVRKKKYHLATFHRGVLQESKQGSTKDPDCTFVSFTPDPSIFPEFT

FNHDFLKDKIRQYTYLHSGLEIRFNDEVFISHNGLKDLFDAEITEPPLYSELFFQNEDLT FNHDFLKDKIRQYTYLHSGLEIRFNDEVFISHNGLKDLFDAEITEPPLYSPLFFQNEDLT FIFSHLEGNTERYFSFVNGQETLDGGTHLTAFKEAIVKGVNEFFGKFFVSNDIREGIVGC IAIKIASPIFESQTKNKLGNTQIRSSLIKDVKEAIVQALRKDKVAPELLLEKIKFNEKTR KNIQFIKQDLKSKQKKVHYKIPKLRDCKFHYNDRSLYGEASSIFLTEGESASASILASRN PLTQAVFSLRGKPMNVFSLEETKMYKNDELFYLATALGITQNEIGHLRYNKVILATDADV DGMHIRNLLITFFLKTLLPLVENNHLFILETPLFKVRNKTTTLYYYSEQEKMQALQQFGK KDSSLEITRFKGLGEISPKEFAAFIGPEIRLTPVTITSLESISSILQFYMGKNTKERKQF IMDNLITOF

CPn_0716 803466 804902
gyra-DNA Gyrase Subunit A
FMRDVSELFRTHFMHYASYVILERAIPHILDGLKPVQRRLLWTLFLMDDGKMHKVANIAG
RTMALHEHGDAPIVEALVVLANKGYLIDTQGNFGMPLTGDPHAAARYIEARLSPLARETL
FNTDLIAFHDSYDGREKEPDILPAKLPVLLLHGVDGIAVGMTTKIF PHNFAELLKAQIAI
LNDKKFTVPFDFPSGALMOPSEYQDGLGSITLRASIDIINDKTLVVKGICPGSTTETLIR
SIENAAKRGTIKIDTIQDFSTOVPHIEIKLPKGSRAKEMLPLLFEHTEGQVILYSKPTVI YENKPYECSISEILKLHTTALQGYLEKELLLLQEQLTLDHYHKTLEYIFIKHKLYDSVRE VLAINKKISADDLHQAVLHALEPWLHELATPYTKQDTSQLASLTIKKILCFNEEACTKEL LAIEKKQAAIQKDLGRIKEYTVKYLKGLLERHGHLGEPKTQITNFKTAKTSILKQQTLI

804968 805306

CTT556 hypothetical protein
IRIKFIDTITIVVRMEPRHIYIRKPETPKAPDVEKPGVPEYMTMANTPTFEGPVKTLDQL
RRALIEQRGAEEGQNMYDNFIQSILISTFGLVHKDMDPAQKASKRMRSVYKEQ

805300 805626

CTT-57 hypothetical protein
RAYMSETYFLALPVDRLMQERFLCSPKRWAPFINSPLYLTLIADHDTPYLAKNLDKFPLP
VEYWEKTYLHYSSLLKSIFLCSDLSSLRLLACTKFEILTLHDLYCAQNI

R05877 806890

CPn_0719 865877 806890

strib (Proudom Gine Synthage)

PDIFFERENT BETTE STREET

PDICKLICKEGRTIKAIR

807671 808489

KGSA-KUN SYNENELSSE KRMMFNNKMILIAGECVIEGEDITLEIAGKLQSILAPYSDRIQWFFKSSYDKANRSSLN GFROPGLTEGLRILAKVKETFGVGILTDVHTPODAYAAAEVCNILQYPAFLCROTDLLVA TAETMAIVNLKKOOPISPWDMEGPIHW/LSTSBRIKILLTERGCSFGYBBLVSDMRSIPVL GROSEFVIERATUSVALKSALSTBRIKSTLUBERT GLURAALAAVSAKGS EETBETNEFIAKS DAASMLSLEEPAALLUWUNGETPISSERMISA

308974 808477 CPn .0722

CFN_U722
CT654 hypothetical protein
YGLSMTKFLYCGLFYSLGLLVLAFGTMVAIIQVDQICDVSCMNKHFQESPPFLKIKKVNV
SKQICSPEERFHCKINGSCMELHFPQSSYSCKEYLTRISGHILTQNFEKQMQFRGNSGL
LNYQDGSLHVYDCRFQVDPVPGYGSPDKEDSSSGGMKTLYLSLFRN

CPn_0723 808978 809703
yhbg-ABC Transporter ATPase
ASMPILSVCNLVKKYNKKPVTNDVSFQINPGEIVGLLGPNGAGKTTAFYLTVGLIRPDSG
KIIFKNVDVTKKTMDKRARLGIGYLAQEPTIFKELTVQDNLICILEIIYKARKQQSHLLN
TLVDDLQLGSCHKKAGTLSGGERRRLEIACVLALNPSVLLLDEPFANVDPLVIQNVKYL
IKILAGRGIGILITDHNAKELLSIADRCYLIIDGKIFFEGSSSQMISNPMVKQHYLGDSF 809703

CPn_0724 810602 809706

No robust homolog present in Genebank/EMBL as of 11/7/98

RTSTRLDYRSGCILSKILPFPELWKMLLGFLCDCPCASWCCAAVANCYDSVFMSRPEHKP
NIPYITKATRRGLRMKTLAYLASLKDARQLAYDFLKDPGSLARLAKALIAPKEALQEGNL
FFYGCSNIEDILEEMRPHRILLLGFSYCOKPKACPEGRFNDACRYDPSHPTCASCSIGT
MMRLNARRYTTVIIPTFIDIAKHLHTLKKRYPGYQILFAVTACELSLKMFGDYASVMNLK
GVGIRLTGRICNTFKAFKLAERGVKPGVTILEEDGFEVLARILTEYSSAPFPRDFCEIH

CPn_0725 810829 810587
CT652.1 hypothetical protein
SCGDVGMFFAPLLYESLRRGLMHPTSHMQQQLARLEFINDQLTTELEHVNELLCSLGFPE
GLTTIKAIAEEVLSDDEPLLD

813384 810880

CPI_0727
813559 816192
CPI_0727
813559 816192
CPi_0727
813559 816192
CPi_0727
813559 816192
CPi_0727
813559 816192
CPi_0727
813559 816192
CPi_0727
813559
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359
81359

CPn_0728 818483 816525
CHLDN 76kDa Homolog (CT622)
VFMNNPIGPOPIDETERTPPADLSAQGLEASAANKSAEAQRIAGAEAKPKESKTDSVERW
SILRSAVNALMSLADKLGIASSNSSSSTSRSADVDSTTATA/PPPPFTFDDYKTOAOTAY
DTIFTSTSLADIQAALVSLODAVTNIKDTAATDEETAIAA/WETKNADAVKVGAQITELA
KYASDNOA ILDSLGKLTSFDLLQAALLQSVANNNKAAELIKEMQONPVVPCKTPAIAQSL
VDQTDATATOIEKGGNAIRDAYFAGONASGAVENAKSNSISINDSAKAAIATAKTOIAE
AQKKFPDSPILQEAEQMVIQAEKDLKNIKPADGSDVPNGTTVCGSKQQGSSIGSIRVSM
LLDDAENETASILMSGFROMIHMFNTENPDSQAAQQELAAQARAAKAAGDDSAAAALADA
QKALEAALGKACQQQGILNALQGIASAAVVSACVPPAASSIGSSVKQLYKTSKSTGSDY
KTOISAGQVDAYKSINDAYGRARNDATROVINNVSTALTRSVPRARTEARQPERTDOALAR
RVIGGNSRTLGDVYSQVSALQSVMQIIQSNPQANWEEIRQKLTSAVTKPPQFGYPYVQLS
NDSTQKFIAKLESLFAEGSRTAAEIKALSFETNYLFIQQVLVNIGSLYSGYLQ 818483 816525

CPn_0729 819905 CHLPN 76kDa Homolog (CT623) CHLPN 70KD3 HOMOTOT (CT023)
PAWSOVETUNI DTKDTMKKQVYQWLASVVLLALTISGYAELFLGEQKVKSHTYTTLDEVK
DYLSKKGYVETIKQDOVLR LAGDVRARMLYFRED LKIFEDROKYNPLPVNRYRSEFYLYI
DYRAERNWLSSKMMMTA LAKSENTAAGVY INRAFLGYRFYKNPETRTDFFME IGRSGLGD DYRAGENMUSSRINMTATAGGENTAAWY INKAPLOTREY KRIPETKTOF FMETGROGGOD LPEGEVOPOINPECLEH YWRTELISKOPLAWFYKY CYWWILDGKHEGVUW INGGKRELYLY GAPEMHILAKATETILMGKELLAWFYKATTLG LRKAGIMGATYRKEYVEALGVPETDVG GRIGHLIKHEPHAQATAANY DEKEAN FETTYKOFSAL YMYG ITTGLGFRAYGAYGKPANDK LGGDPTEKKEDIGI I GAP

CDT_0730 821676 819963 mvin Indestrat Membrany Protein CCFKCZGENGIMERKIAHEVSTARSIFNILEXZPPCSEPTGIFREIAMATYFSADDIVAAFW

LISERTVEFLERILUS LILL HEEFLRAGSERAAFERREGRLIKGST I EFTLLIE AVLAVVLOYVEEGYEMILL LLEGGIFLMANNOALLHGENKEFFYSLAPVVNII WIFFVIAARHSDPRER I IGLSVALVIGEFFEYL ITVEGVWKFLLEAKSPEGEHDSVRALL APLSIG ILTSSIFOLIHLSDIGLARYVHEIGFLYLMYSLKIYQLPI HLEGFGVFTVLLPA ISRCVOREDHERGIKLMKEVLTLTMSVMI IMTAGLLLLAL POURVLYEHGLEFOSAVY URVLRGYGASI I PMALAPUSVLEYYAGROYAVPLE IGIGTALANIVLSLVLGRWYLKDVS GIGYATSITAWYQLYFLWYYSSKRLPMY KKLWESIRRSIKVMGTTMLACMITLGLNILT OTTYVIELNELI AMPLGSITAOAIAYLGESCIFLAFLEFGFAKLLRVEDLINLASFEYW BEREGLEFOR OTTOL 5-4-11.13.40.117M-

CPn_0731 821474 821760 No robust homolog present in Genebank/EMBL as of 11/7/98 VALAISRNIPVIRLOM/PDNILK/ERAKETSLSFLLIKPFSPPPLKQDYLFDISPYTSSE ITIGGSYFKLNKASLQSSTLRLKSISIIS

822/092 CPn 0732 CPH_0732 82909 622976

nfo-Endonuclease IV

nfmkvLpppSipLLGahtSragGlknaiyegrdigastvQiftanQrqwQrralkeevie

DfkaalketDLSyimShagylinpGappDylLekSriGivQeildcitLdisfvnhfpG

AlkSskeDCmyRivSsFgOSaPLFDSSPPLVVLLETTAGQOTLIGSNFEELGYLVQNLKN

QIPIGVCVDTCHIFAAGYDITSPQGWEDVLNEFDEYVGLSYLRAFHLNDSMFPLGANKDR

HAPLGEGYIGKESFKWLMTDERTRKIPKYLETPGGPENWQKEIGELLKFSKNRDS

823739 823101 CPn_0733

rs4-S4 Ribosomal Protein
GLKYMARYGGPKWRVARRFGANIFGRSRNPLLKKPHPPGQHGMQRKKKSDYGLQLEEKQK LKACYGMIMEKGLVKAFKEVIHKOGNVAOMFLERFECRLDMMVYRMGFAKTIFAAQQLVA HGHILVNGRRYDRRSFFLRPGMQISLKEKSKRLQSVKDALESKDESSLPSYISLDKTGFK GELLVSPEQDQIEAQLPLPINISVVCEFLSHRT

823863 824915 CPn 0734 ycea Ontkehfsskgnflocnyfodyvrvfimekkyyalayyyitrvdnpheeialhkkfledl Dvscryviseogingofsgyephaelymowlkerpnfskikfkhhikeniffritvkyr Kelaalccevdlskoakhispoewhekloenrclildvrnnyewkichfdnatlpdigtf Reffeyaeklagecdpettfvmmyctggircelyspvllekgfkevyoldogviayogov Gtokwloklfveddrlaifidesdpdvafiaecchcqtpsdayyncantdcnalflocde Cimohogccgeecsgsprvrkfdssrgnkpfrrahlceisensesasccli

Pn_0735 825680 825003

*Uridine Kinase (Uridine Monophosphokinase) (Pyrimidine Ribonucleoside Kinase).

GEKFMLMMLMMIGITGGSGACKTTLTONIKEIFGEDVSVICQDNYYKDRSHYTPEERAN LIWHHPDAFDDLLISDIKRIKNNEIVQAPVFDFVLGNRSKTEIETIYPSKVILVEGILV FENQELRDLMDIRIFVDTDADERILRRMVRDVQEQGDSVDCIMSRYLSMVKPMHEKFIEP TRKYADIIVHGNYRQNVVTNILSQKIKNHLENALESDETYYMVNSK

CPn_0736 827731 825992
ygeD-Efflux Protein
RGELLKLARQCLVAFMTVSVKKKSFRALVTTHFLTIINDNLYKFLLAFFLLEGKTLTENA
KILSCVSFFFALPFLLLAPLAGSLADRFQKRNIILATRFIEILCTILGTYFFFIQSVVGG
YVVLILMACHTTIFGPAKLGILPEMLPSEQLSQANG INTAATYTGSILGSCLAPLLVDVT
HRLGVNSYWPTIMCVIVSIISTLISFGIRPSNVMVQKITLVSFKDLWKVLKDTRNIH
YLTVSIFLGSFFLLIGAYTQLEIIPFVEFTLKYPKHYGAYLFPIVALGVGTGSYITCKIS
GKDIKIGYVPLAAIGLALVFMGLYAFACSILFVLFFLLALGFLGGVYQVPLHAYVQYASP
EHKRGQILAANNFLDFFGYLVAAGVIRVLGSNLGLSPETSFFYIGWFVLAVSINTLMIWR
EHVYRLLLGIILRRQLGYYLKHQSSSPKCYFVAVQSYREIRRVLAALTKTVRSRVIILD
QKLVPGWRAWLLSWCVPTVVSSVRDNDSEAQDAWAVLQANHLKTSLKKFPDVSVVCLGLP
KNYERTFISIQEGGIDLHPIOLVOREGKRVIYTLVFPHA 825992 827731 KNVERFTSILQEQGIDLHPIQLVQKEGKKRVIYTLVFPHA

CPn_0737 827469 830756

"recC-Exodeoxyribonuclease V, Gamma"

KRSAKLPASGASKRKGRAKKKLTQERIFAFSVRVLPSNRKNAKRNLYKLSFIIVRKCVVT
SALNDFFLTETVMNATKHCRASFSNSPRHLLAQLAEDITSTHQKPFTRAWILVANATTGH
WIKNQLVMVLSDHIFMGSTIFTASDSIVKHLFLGSGSGOPNIPDDVLTLPLLINNILEEIS
KASKFENGREFLSPFTYETTKKLAAAFKQFHTFSQRPTKNASHYQELFQILESHFSSYEE
MFTTILNNRTQEEDCSLHIFGYAHLPKHLAEFFINLSTYFPVYFYCFSFCEYFGDLLSK
HDSSLGVIQNSILDLKPTSPODFSCTKGTICIYRALNIFREVQEVFCKVTELHARGVSPE
EIFILSSHIESYKVHLNAIFFNEHVPIYYFDEVDPRAEDLRNKILLLSSILQTQGDLHYIL
OLLTHPQLQQPIDONKVPYLIKKLSSENGKISSKDRASGQOMKALGDLILEEYPFHQEGG
RVSQVEVWETTVPLIYFIQERINLYLSSSQHSYEDLFQNVTSCLEKIFYLSPETSFITL
LGRNKTTSSDIFDLLNRTTTHEELAFSSTEDEENFHFLQILVSTKHELHISYISSAQFN
LGSNKTTSSDIFDLLNRTTTHEELAFSSTEDEENFHFLQILVSTKHELHISYISSAQFN
LDSPFLANHIKETLOLPVETLEFTQPYLSAFFKNKACLHTSGEVTNYSLAHAFYSKKALLPSL
FIPTVKQVNLPQHLSLNEIIKGIFSPLDLFLKTTNNLRISYPEHLKKQCKLFFTKHQIED
FWNECFVDKEHDLIPSISPHABELFTYYREKTILLRNGLDKDPKHSPYTYTSSSIFED
ENHECTVDKEHDLIPSISPHABELFTYYREKTILLRNGLDKDFKHSPYTYTSSSIFED
SQQFIPLLSPLCWRTUDDEEKFHQAVLSAISEEAKNPSLPIFWQFHNRNIEEILNNVGAS
ERLKILSLFRGPCEAV ERLKILSLFRGPCEAV

CPn_0738 830719 833895

"recB-Exodoxyribonuclease V, Beta"

Figh-Exodoxyribonuclease V, Beta"

AITEPNASTNELKVRIKDNLSOTKEFLEASATTCKTFTIEQIVLRALIEGSLTHVEHAL AITEPNASTNELKVRIKDNLSOTKEFLEASATTCKTFTIEQIVLRALIEGSLTHVEHAL AITEPNASTNELKVRIKDNLSOTKERLKAVLNSOPASLPTYLDINCNVKQIYMQVRNALA TLDOMSLETIRGENFYLEOYPERTELHKNPALTHSQUVLHHITNYLKQDLWKNVLFOE OFHLLAVRYNTTSHKTSLTVKLKLASYTOPIGSYFGSKYPELEQIGLMHQVIYNSLEIP KOVFLOQUTAHISGFKKOPFSILDDHHFVDLLYTSETHSGLFSFFKIAETPNFKHRLAR YKRCAAPTVLENMENDERTEPCHIDGRIFFTLLYDDLOGYLKONYTTPMLSPDESVFALEKL LSSCSAQPOVQALERQYQLVLIDEPOIDTOKOWNG FFOILETISFKTGSSLFLIODPKQSIY EMRJADLIPTYLTAKSSESEDKOLGLVINNYRTTPHLMBALIDGIFGKISFLEIPGVLPIESY HALHIPOSETFENDHAR LIBEFTFTI KODALAHIFSELEAFOLGVKEKKI FINITETTI KOTALAHIFSELEAFOLGVKEKKI FINITETTI KOTALAHIFSELEAFOLGVKEKKI FINITETTI KOTALAHIFSELEAFOLGVKEKKI HITTI HISTORIETI TI HIDISAGLEYDI VPČEG LEKCKKNRGSSELLISEMTVA TPAPAVOLIVLI LOTOPPSLORDSALTNYKLEGTOSSAYD LATHLIGGENDLEFSSLEKDISTI KODALAHIFTALLETTALKYTPPKT LESPSTKFLLDTUK LATHLIGGENDLEFSTKFLLDTUK LEGOLPPOLLOGNSALTNYKLEGTOSSAYD LATHLIGGENDLEFSSLEKDISHATTYN NILPLETFALKYTPPKT LESPSTKFLLDTUK LATHLIGGENDLEFSTKFLLDTUK LEGOLPPOLLOGNSALTNYKLEGTOSSAYD LATHLIGGENDLEFSSLEKDISHATTYN NILPLETFALKYTPPKT LESPSTKFLLDTUK LATHLIGGENDLEFSTKFLLDTUK LEGOLPPOLLOGNSALTNYKLEGTOSSAYD LATHLIGGENDLEFSSLEKDISHATTYN NILPLETFALKYTPPKT LESPSTKFLLDTUK LATHLIGGENDLEFSTKFLLDTUK LEGOLPPOLLOGNEYTMETTIRF EKHTILLEG EPETLIKLLSKTFFSSLEFSCONSCHOLTTY INGELEDOPSCLOGNEWYN MET TIRF EKHTILLEG COPCILLOGNEYTMETTIRF EKHTILLEG COPCILLOGNEYTMETTIRF EKHTILLEG COPCILLOGNEYTMETTIRF EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCILLOGNEYTMETTALKTERD EKHTILLEG COPCIL 830719 833895

CV LE LEC LUTQUINGEFALNOCIED L'ENFINERA LUKCOA

R34892 CPn_0739 833861 CT 368 hypothetical protein CKVLFKLMSYSLRNKYTKICVYIIIALGILSFRSIPQEVYDKIRSSFVSLHVKFFPKIKQ

APSSHLANLELENLVLKERVASLEEKLKLYEVSNHTPPLFPEILTPYFHKLVEGKVVYRD YTHWSSSCWVNVGKTHGIKKNSPVLSGNVLVGLVDYVGEHQSRIRLITDVGMKPSVVAMR OD LOGIMMENGUREL POVEO I SHAY TLEKONYEKT SOLOELDS LTOGEGENOALLRGT L POVE PALMYE, DIE TEKNIK REPERINGKTELLENE LYTTOLD TVEDEGLEVARVTKVEAERO PACTER LEZA, OLGEKEMEGLEJER ELEPTERTEHERE TEGLEND

836054 834864 CPn_0740 LPT_U/4U 833634 834864
LYTB-Aromatic AA Aminotransferase
SYMSFFNHIPTFSPDAILGLQNVFFADKRPEKVNLVIGVYEHPQKRYGGLSCIRKAQTVI
LEEEQNKSYLPISGLQIFLDEWRELVFGAVDPSAIVGFQSLGGTGALHLGARLLSVAKGS
GKVYVPPGVMSNHIRTFSQGGLEVIRYPYYSKEQKQLLFEPLIAFIKEVEKNSVILLIGC
CHNPTGVDFTEDMWKELAILMKERELIPFFDTAYQGFAHGIELDRKPIEIFISEGNTVLV AASSSKNFALYGERVGYFAVHSTFTDELVKIHSFLEEXIRGEYSSPQRWGVEIVSTILSN PYLKEEWQSELNFIRESLGKMRTRFVQALRKVAGHTFDFLLSQHGFFAYPGFSDKQVLFL REQHAVYTTAGGRMNLNGITEKNIDHVVQSFIQAYEL

CPn_0741 838383 836185
greA-Transcription Elongation Factor
EYIFRLKTGDIVDYLEKLOVLIEEGGSANFLSLWEEYCFNDVVRGRELVEILEKVKSSSL
ASLFGKIVDTVVPLWEKIPEGKDKDRVLQLILDLQTSNSQMFFDIATEYVNKKYSGENF
NEALRVVGLRDGRDFGFSLSRFDFLMHMHKGNFVFNQGGWCVEVMGVSFLQQKVLIEFE
GIMSAKDISFETAFKSLTPLSGDHFLSRRFGDPDGFEAFAKENPIEVVEILLRDLGPKTA
KEIKDELVDLVIPEADWNRWMQSAKTKIKKGTRIISPDNPKEPYVLSDACCSHMGQLERK
LGLSLNSAEKISLIYHFIRDLHSELKNIEIRKSLVKALQDLDVEEGNKSLILQRELLLSE
YLGIKDASIDKEYITSLSEDDTSRLLENMPIVALQKSFLSLVRKYSSFWQQVFMQILLYT
TSPTMRDFVYKTIKNDPSSVEVLKKRLLDSAHQPMMFPELFVWFFLKLGNHEDGLFDPED
KEVLRLFLESALNFMYQVASTPHKELGKKLHHYLVGQRYLAVRQMIEGSSLPFLKELLLL
STKCPQFSSSDLNVLQSLAEVVQPTLKKHKSNVEEENVLWSTSESFSRMKAKLQSLVGKE
MVDNAKEIEDARSLGDLRENSEYKFALEKRARLQEEIRVLSEEINRARILTKDLVFTDKV GVGCKVTLKGDAGEVVEYTILGPWDADPDSCILSLQSKLAQNMLGKKLNDVVILQGKEYK

CPn_0742 838442 838888
CT635 hypothetical protein
TKMMVIVMNSKSAOKIIDSIKOILTIYNIDFDPSFGSSLSSDSDADYEYLITKTQEKIQE
LDKRAQEILTGTGMSKEOMEVFANNPDNFSPEEMLALEKVRSSCDEYRKETENLINEITL DEHPTKESKRPKOKLSSTKKNKKKNWIPL

CPA-0743 838956 840362
"ngka-Ubiquinone Oxidoreductase, Alpha"
IPHKITVARGLDLSLOGSPKESGFYNKIDPEFVSIDLRPFOPLSLKLKVEQGDAVCSGAP
IPHKITVARGLDLSLOGSPKESGFYNKIDPEFVSIDLRPFOPLSLKLKVEQGDAVCSGAP
IAFKKHFPNTYITSHVSGVVTAIRRGNKRSLLDVIIKKTPGPTSTEYTYDLOTLSRSDLS
EFFENCLFALIKORPFDIPAIPTQTPRDVFINLADNRPFTPSPEKHLALFSSREEGFYV
FVARVRAIAKLFGLRPHIVFFDRLTLPTQELKTIAHLHTVSGPFPSGSPSIHIHSVAPIT
NEKEVYFTLSFODVLTIGHLFLKGRILHEQVTALAGTALKSSLRSVITTKGASFSSLIN
LNDLSDNDTLISGOPLTGRLCKKEEEPFLGFRDHSISVLHNPTKRELFSFLRIGFKNFT
TKTYLSGFFKKKRTYTNPDTNLHGETRPIIDTDIYDKVMPMRIPVVPLIKAVITKNFDLA
NELSFLEVCGEDFALPTLIDPSKTEMLTIVKESLIEYAKESGILTPHQD

CPN 0744 841387 840389
hemB-Porphobilinogen Synthase
EMSSLTLSRRPRRNRKTAAIRDLLAETHLSPKDLIAPFFVKYGNNIKEEIPSLPGVFRWS
EMSSLTLSRRPRRNRKTAAIRDLLAETHLSPKDLIAPFFVKYGNNIKEEIPSLPGVFRWS
LDLLLKEIERLCTYGLRAVMLFPITDDLKDAYGSYSSNPKNILCHSIHEIKNAFPHLCL
ISDTALDPYTTHGHDGIFLNGEVLNDESVRIFGNIATLHAEMGADIVAPSDMMDGRIGYI
RSKLDQSGYSKTSIMSYSVKYASCLYSPFROALSSHVTSGDKKQYQMNPKNVLEALLESS
LDEEEGADILMYKPAGLYLDVIYRIRQNTCLPLAAYQVSGEYAMILSAFQQGWLDKETLF
HESLITAIKRAGADMIISYSAPFILELLHQGFEF

CPh 0745 841903 841742
No robust homolog present in Genebank/EMBL as of 11/7/98
VDSCFDDWRASSLOGSTTYNVAYDPKHTLAYGFCNQVSVKKFHLKPPKSQEKFL

CPALE 746 841939 843567
CT632 hypothetical protein
FSGRCPFSFEVFMLGKEEFTCKQKQCLSHFVTNLTSDVFALKNLPEVVKGALESKYSRS
VLGRALLLKEFLSNEEDGDVCDEAYDFETDVQKAADFYQRVLDNFGDDSVGELGGAHLA
MENVSILAAKVLEDAR HGGSPLEKSTRYVYFDQKVRGEYLYYRDPILMTSAFKDMFLGTC
DFLFDTYSALIPQVRAYFEKLYPKDSKTPASAYATSLRAKVLDC IRGLLPAYTLTNLGFF
GNGRFWQNLIHKLQGHNLAELRRLGDESLTELMKVIPSFVSRAEPHHHHHQAMMQVRRAL
KEQLKGLAEQATFSEEMSSPSVQLVYGDPDGIYKVAAGFLEPYSNRSLYDDLIDYCKKMP
HEDLVQILESGVSARENRRHKSPRGLECVEFGFDILADFGAYRDLQRHRILTOERQLLST
HHGYNFFVELDTPMEKSYREAMERANETYNEIVQEFPEEAQYMVPMAYNIRWFFHVNAR
ALQWICELRSQPQGHQNYRTIATGLVREVVKFNPMYELFFKFVDYSDIDLGRLNQEMRKE
PTT

843949 844053 CFN_U747 843949 844033 CT631 hypothetical protein RTCMGCKGAEVQILSSRSLSGMKILSSSLFYKKFC

844996 CPn_0748 R44996 844121
ispA-Geranyl Transtransferase
GTLVLHALDTYRPSIESALEKALEGFGPIGHPIRSPVEYALOGGCKRLRPGLVCMMAQGL
GLANIDVMDSALAVEFVHTOTLIADDLPCMDNDDERRGRRPTVMKAFDEATALLASYALIPA
AYSHLELNAKKLKEQGCDPREIDIAKNIIGGTGWLGGGYDDMFFSNRGQEHV
QSIMIKKNGSLFEIACISGWLFGGGDPGFAPIITSFSNAFGLLFQIKDDFSDLCKDSQOI
GLAYALLFGEKAALELLARSQNNGLELLDRLSANGLKWDSEFETIISSLGSF

CPn_0749 84563H 845006 glmU-UDP-clicNAc Tyrophosphorylase GHID ODE A COME. LY CONDUCTION CASE
VCMMYTIANS LECHEDELY HELLER KILLY UND LLD MUDMLENHVERG HIGTVECGVTLKN
LEK LE LAEDAYVER AY LYCHOLIC GOVEVERIG AY LOGNY TROSRC VACHOTTETKNSYLG
HITKAMHEAY IN DOVLESIEVNLGAGVROANER (DEPHLY VROTERIKSK LDTORRKLGAF
LAKOVA I CENYV I NI VOH LLDHTR LEDGOV I EMDOMLENHVESG HIGTVESGVTLKN

Clar 0750 846465 845/07 totD/cpxR-HTH Tran

onal Regulatory Protein - Receiver

Doman KITDFILRIHSYNLFCFHMIGDKIILFVTEDLSI SSQLKDLASQRSDYQILVSPVFPTSF ESVAIFCEYLLIPEDIFSPGIFSPEDLIVLFDTFOEEAITKVLNOGATGYLLRFITAKVL DAVIRAFLRGHEVLEHSIPDTMTFGDHTFRVLMLVIESPEGSVYLTPSEAGILKKLLINR GHLCLRKNLLAEIKGNTKEIIARNVDVHIASYRKKLGPYGSKIVTIRGVGYLFSDDDSIP LONHONTAHPNEE

CTBLEVEL MAKERS AT LA FREY CTBLEVEL CTBLEVEL MAKERS AT LA FREY CTBLEVEL MAKERS AT LA FREY CTBLEVEL MAKERS AND A CTBLEVEL MAKERS AND YLYSMWLTKVAPSPQSYRLAGLCLMENKRYDEALEFLCMLSPNDSINDYKTQKALAFCQK HQSKDRAAS

CPn_0752 48595 850082

reco-Exodeoxyribønuclease V, Alpha
GWALHTEFAPFLEDLYHÖOVISFLDIAFASKHISSDFEESFYFLAVSSALWRYGHPFLSL
EENRIRPSLOGISETDLYRGFHNLPKEARDKLFVVVSGRLYLRSLYTIRSKLLDKLSLLC
SATPNYFPPSIDSSILSEEGONFIFNKITOGCFSIVSGGPGTGKTFLAAQLILSLVKQOPK
KRIAIVSPTGKATSHIRQILHKYNIFDDMYLMGTVHHFLQEYAYRRYNSIDVLLVDBGSM
VTFDLLYSLVOTLOGYEKDKKLYTSSLIILGDTNQLPPIGIGVGNPLQDLIGVFHENTFF
LKTSHRAKTGVVDÖLTQSVLRGEMISFSPLPSISSAIEVLKNRFVXSLRQSEARLCVLTP
MRHGFWGVLNLMYMHQRLARSDPDLRIPIMVTSRYETWGLFMGDTGLLCLKTQKLHFPQ
HEPIDSRALSOJVYNYVMSVHKSQGSEYDEVIVIIPKGSEVFGVSILYTAITRAKYRVSV
WGDPETLHKIJKKSNY

CPn_0753 851009 850161
No robust homolog present in Genebank/EMBL as of 11/7/98
IMATAHLØRQALLHLRSWTPAIRASGNLFRQSMSLHNNVLFAGDIVGAIKNSTAISRHA
LGSSHYÆHAALQKTEGFIGAADGVWTAVAGAHLWGQLLHGSMIFETDEETGELRCNEAD
AEGCMTÖKLORRSALT ITGKVARLASKTICTATFLHEMDVVSLGANANKIGCKVTSCLNL
VATGCELTESSISLYRILSTRPETISDPENRNKPSAEFAARSKAIRNAFIAWLGDVVDLV
CDALØTLSLFLPAILGVHAVLIMAILGLISCVINFVKDYAKIG

CPn_0754 851381 851040 TSQU-SSU RIDOSOMA1 Protein
OFILNLKVLVLSGDIMAPKKPNKONVIQRRPSAEKRILTAQKRELINHSFKSKVKTIVKK
FEASLKLDDTQATLSNLQSVYSVVDKAVKRGIFKDNKAARIKSKATLKVNARAS

851579 (CPL_0755 851579 852799
TT616 hypothetical protein
YKDLFFMLLVRKWLHTCFKYWIYFLPVVTLLLPLVCYPFLSISOKIYGYFVFTTISSLGW
FFALRRRENOLKTAAVOLLOYKIRKLTENNEGLROIRESLKEHOOESAQLQIOSOKIKNS
LFHLOGLLVKTKGEGOKLETLLLHATTEENRCLKHOVDSLIGEGEKTEEVOTLNRELAET
LAYOQALNDEYOATFSEORNWLDKRQIYIGKLENKVODLMYEIRILLOLESDIAENIPSO
HSNAVTGNISLOLSSELKKIAFKAENIEAASSLTASRYLHTDTSVHNYSLECROLFDSLR
JEENLGMLFVYARQSORAVFANALFKTWTGYCAEDFLKFGSDIVISGGKOMMEDLHSSREE
CSGRLVIKTKSRGHLPFRYCLMALNKGPLCYHVLGVLYPLHKEVLQS

CPI_0756 852889 854676

ITOD-RNA POlymerase Sigma-66
ISYLPLIKLSSKARNPLVLFOVRKLFMNTONSQATEVSSEEESQKKLEELVALAKEQGFI
TYEEINEILPMSFDTPEOIDQVLIFLTGMDIQVLNQIDVERQKEKKKEAKELEGLARRTE
GTPDDPVRMYLKEMGTVPLLTREEEVEISKRIEKAQVQIERIILRERYSAKEAISIAHYL
ISGKERFDKIISEKEVEDKTHFLKLLPKLITLLKEEDTYLENLLLSLKQPDLSKQEAKL
NDSLEKCRIRTOAYLRCFHCRHNVTEDFGEVVFKAYDSFLHLEQQINDLKVRAERNKFAA
AKLAAAKRKLYKREVAAGRTLEEFKKDVRMLQRWMDKSQEAKKEMVESNLRLVISIAKKY
TNRGLSFLDLIQEGNMGLMKAVEKFEYRRGYRFSTYATWWIRQAVTRAIADQARTIRJV
HMIETINKVLRGAKKLMMETCKEPTPEELAEELGLTPDRVREIYKIAQHPISLQAEVGEG
SESSFGDFLEDTAVESPAEATGYSMLKDKMKEVLKTLTDRERFVLIHRFGLLDGKFKTLE
EVGSAFNVTRERIRQIEAKALRKMRHPIRSKQLRAFLDLLEEEKTGTSKVKSLKSK

CPn_0757 854709 855134
folx-Dihydroneopterin Aldolase
PCIKNIALVIAIERYOLIISKFRMMLFLGCSVEERHFKQPVLISVTFSYNEVPSACLSDK
LSDACCYLEVTSLIEEIANTKPYALIEHLANELFDSLVISFGDKASKIDLEVEKERPPVP NLLNPIKFTISKELCPSPVLSA

CPn_0758 855104 956459
folP/dnps-Dihydropteroate Synthase
RAMSEPREVCLSLGSNLGNREKNLQIARTLLGEQAVLGLRSSVILETEALLLPGSPPEWD
LPYFNSVLVGETTLSLRELLVTIKQIEKVCRAEESPEWSPRTIDDVILLYGDESFCCDH
LPYFNSVLVGETTLSLRELLVTIKQIEKVCRAEESPEWSPRTIDDVILLYGDESFCCDH
TEITIPLSNLLSRFFLIALIASLGPYRRFCTGOSPYHNFTFGELAHHLPSPPGMIRRSLS
PDTMLMGVVNVTNDEMSDCGMFLDPEKAVAQAEKLFTEGAAVIDFGAQATNPKVKQFLSV
DQEWERLEPVLRLLKETWSNRKQYPIISLDTFYPEIILRAMDIYPIQWINDVSGSGSMFA
EVARDCELSLVMNHSSSLPVDPKNILSFSVPIGEOLLSWGEQKLMFSDVGLNANQVIFD
PGIGFGKGAAQSLATLVEIAKFKRLGCPILIGHSRKSFLSLFGNHDPKDRDWETVGLSIL
LQQQGVDYLRVHNVAAHQKALSVAACEACAPI

856434 CPO_0759 8504.4 30000 fola-01hydrofolate Reductase LLVKPVIPCNEENPLSVEDERNPSVEGTVACDERGVTGLEGKLPWHYPEDLOFFSETTOK PPTWGRKTWETLEPKYPVDRAVAVECHEKROSVHGETWYPSLEGEFLLLDLSSPTFLTGG GELYGLFLENQTVPDFFTSHTKKEYA/DPFPLSLLETWTKTVLRDTQKTTTCYYENHHS 356997

CPn 0760 886063 CHOLOTOR SSORTS STANDARD PROPERTY STANDARD COMMENTAL ACTION OF STANDARD COMMENTAL PROPERTY OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENTS OF STANDARD COMMENT

857698 858375 CPn 0761

CFN_07AL 85768 858375
CTh10 hypothetical protein
GIMTSWTELLDKOIEDOHMLKHEFYORWSECKLEKQQLQAYAKDYYLHIKAFPCYLSALH
ARCODLQIRRQILENLMDEEAGMPNHIDLWRQFALSLGVSEELLANHEFSQAAQDMVATF
RRUCDMPQLAVGLGALYTYFEIQIPQCVEKIRGLKEYFGVSARGYAYFTVHQEADIKHAS
EEKEMLQTLVGRENPDAVLQGSQEVLDTLWNFLSSFINSTEPCSCK

5 M. 16 /2

100 A ROBA DOMESTICATION DESCRIPTION OF STREET SERVICES OF STREET SERV ATHEISTIKTGALSLOLALGIKOVPKGRVIETFGPESSCKTTLATHIVANAQKKGVAAY
IDAEHALDPSYASLIGVNIDDLMISOPDGEBALSIAELLARSGAVDVIVIDSVAALVPK
SELEGDIGDVNVGLQARMASQALRKLTATLSRSOTCAVFINQIREKIGVSFGNPETTTGG
RALKFYGSIRLDIRRIGSIKGSDNSDIGNRIKVKVAKNKLAPPFRIAEFDILFNBGISSA
GCILDLAVEYNILEKKGSVFNYQEKKLGQGREFVREELKRNRKLFEEIEKRIYDVIAANK TPSVHANETPQEVPAQTVEA

860520 859972 VgfA-Formyltetrahydrofolate Cycloligase
NFPMTDPKIEKSALRKLFISIRRDLSEERKHEASSAVASFVRSFSKESVVLSFVSFNHEI
DMCEANRILIGKCTLALPRIDGENYPVLIPSIDDLISVVHPKDPFSKOTPISSDKITHV
LVPGLAFDQQGYRLGYGHGFYDRWLAQHPYPSIRTIGIGYCEQKIDRLPQESHDIPLSQI

CPn 0764 861819 860524

CT648 hypothetical protein
GYKSMDIKKLFCLFLCSSLIAMSPIYGKTGDYEKLTLTGINIIDRNGLSETICSKEKLKK GYKSMDIKKLFCLFLCSSLIAMSPIYGKTGDVEKLTLTGINIIDRNGLSETICSKEKLKK YTKVDFLAPQPYQKVMRMYKNKRGDNVSCLTAYHTNGQIKQYLECLNNRAYGRYREWHVN GNIKIQAEVIGGIADLHPSAESGWLFDQTTFAYNDEGILEAAIVYEKGLLEGSSVYYHTN GNIWKECPYHKGVPQGKFLTYTSSGKLLKEQNYQQKRHGLSIRYSEDSEDVLAWEEYH EGRLLKAEYLDPQTHEIYATIHEGNGIQAIYGKYAVIETRAFYRGEPYGKVTRFDNSGTQ IVQTYNLLQGAKHGEEFFFYPETGKPKLLLNAWHEGILMGIVKTWYPGGTLESCKELVNNK KSCLLTIYYPEGQIMATEEYDNDLLIKGEYFRPGDRHPYSKIDRGCGTAVFFSSAGTITK KIPYQDGKPLLN

862415 861801

CT647 hypothetical protein
TTIYIKLLGRLMKKWISILILSFLSLLSILPVLAITINHVKISQRWSDLNSQILTLKVIR
DHEDQVIKHNARISKDRNNIS IESLNASCKQLRPLSKERERLINKLNSNSLLAQSKEVWER
KRÆLEKSNHOLVWNCENNISP AFVELEQATEMDNEDIESLFSLFNPENPVAPLVFFTCW
KNTROTTPLGNEVWLTHAEAISRWI

MTROTTPLGNEVWLTHAEAISRWI

CPTC 0766 863785 862394

CTG 6 hypothetical protein

AMTERLPVYHIGLITKAENNTIKIAILOKTCKGWIVCHCEQIPEGKTWSLPKKYFAAPTTF
SLOGSDILVKSSSSSLKNRNILKVALTNILEASLALPWESLIVOPOLGKETDRGETPLTL
WIAGRNTLKKELSFLSOAQIFPDKLSCRAADIFFLAEQSPLKSLPAYLLIYGGSEEVTCI
FVAGHAIAVARSFSNHSTKKSCDDIHATLOYIQETFPOTVLPAIHVAQISPNLOKILEGV
SLELVVCQSNTYGVEDEDWEIYGDTIAAAHHGASRRPLITFPYDATSVSPAAQKHWLLRS
SLELIGKVALMATVVVSLGSVLKLKSLSSSASNHFAFACPEEGVLPRSLKAAEKTVKAIGB
KNSASNYPLLPTIPTSEDTLKFLLALGKSSPSIKFSYFSYTMTSYPSKONPSLPYSALVW
KKOKOPEDIPOFLKKISSHPKLOHVSESLEDORSFKLOFTLSS VKÖÖGOPEDIPOFLKKISSHPKLOHVSESLEDORSFKLOFTLSS

CPh_0767 863878 864177

CTS45 hypothetical protein
NIMLSYLLRTAINVYSFLILAYIFASWYPDCQSARWYQLVSKCVDPFLNFFRRFVPRIG
IDPSPFVGLLCLGILPFVILRVLRFIILNIFHSPWLLQYL

CPTT-0768 864144 865163
yohl/nir3-predicted oxidoreductase
YFSTSMAAPIFIKNILLRSSIVYAPLAGFSDVPYRCMSALYOPGLMFCEMVKVERILYAP
ERTSKLLDYNENMRPIGAQLCGSNPETSGEAAKILEGLGFDLIDLNCGCPTDK/TKDGSG
SGLKTPELIGRILDKIINSVSIPVTVKIRSGWDMEHINVEDTVRIIRDAGAAVFVHGR
TRAGCYHGPSKQEYISRAKAAAGKEFPVFGNGDIFSPEAAQAMLTTGCDGVLWARGTLGA
PWTGKQIQDYLTTGSYEKIFFIKRKAAFLEHMRLVEDYYQSETKFLSETRKLCGHYLISA
AKUPELBSSIAKATSVOFLYOLMNYFFADDGSI AKYRFLRSSLAKATSYQEVYQLVNDYEEADDSSLETFVKC

CPL_0769 867763 865121

topa-dna topoisomerase i-fused to SWI Domain
sigehairlmkksliivespakiktloklugsefyfassighivdipakefgidvdhdf
epoyovlpdkoevinnirklaakcekvylspdpddregealawhianolppspliorysfn
aitknavtealkhprtidmalvnaoparrlldrivgykispilspklogrsgisagryos
valklvvdrekaidafvevetymlrvlmddpktyttfpahlyayockkmekeipeckten
bvllinseekarhyaellekssytitrveakakrrfappffittogeasrhfrfsasr
msiaotlyegovldsedstglitymrtosvrvdpealttvryiootfskeylpekanv
yttkkmtodaheairptdinltpdklknklsddofkvynlipkrfvasoitpahlyotlav
oittdteidlrasgsllkfkgflavyeeroddendeedhflephaodalikeevsoeo
aftkelprfteaslykelekscigrfsytatimnkiosrbyttkenorlpptligkiso
fletnfprimdigftalmedeleliadnkkpwkllloepyttflpvitaekeavipril
tniecskchkoklvximsknsyfyccseypecdyrtsefelafnkedyaedffwdsfcpi cogywkvyndgyoftlecekyfecrtisihkkeeljoepetpepaicorkifkkrsr
nnkifyscseypecsvionsidavitkysotekipykkrtptkkssakttkaaktpskk
ckakssvkkssekktgplflepspolakmignepvsggeatkkiwdyikeholoapenkkl
lyponnlatiigpnpidmfolskhlsghltkvsnessass

868322 869131

CFI_9770 300322 307131 CTT42 hypothetical protein KPRTRNVEKLEFVTSLSSPDDDLITFNKQGL/AGPEEKVAFLVRSNAMLDAGPETPASF PECLEROFO FPEYVEVLY SNEGLOW WEAKETWILLINEST TOLKKHHRKAS RWLGMYSRD
EYLAHEAVHAVRMKFHEPVFEEVLAYOTS BWGWRRFFOPLFRSPGESYLLLFFT LGLGI
SLWYPAG ILLIMLVLTMYFLMRLCMAGSYLYRAMKKIPKMLGV PPLWVLLRLTDKEIKMFA KEL-LANTEHAUKKKTENAKATAOOA

CFn_0771 870511

CPD_0771
POST11 S05111 S0144
Epoll RIA Folymerage Sigma/54
IFYGNCKRLYDSGALDMPQKOKUSLKYLPGURMOGGLOMLOSPLTELSSYVVOEIIDNP
FFDASGAERERWIND YRPYNDTTSYLNOTYSPQECLYTRLPQIERASTAERRIAHQI
AYMIGDDE JULHNEUSPAQELEDILEK THKVWTT UJILDPEG LASPSLOSYMMKLERNSS
HQAYSIVPDCYPLMTNIEFADIMKKFSLOLDELRHILLKKALGSIPWCPAAACTVKPMVS
TPLPGTYLFYSOGGMETEVITYGURSIKLNKETFHFYEHLPKEEQKNLSQOILGAKWLIK

NLRKREOTLLOVMETLLPK GKIPAPYPLSIKDLAEDLSFHESTIFRAIENKAVA APIGIFPLKHLPPRGIHODS ANVLOWIRONIATEGTPLSDSVISDRITAKGIPCAR RTVAKYRAQLKILPANKRKKLFYIRSSNSHFRURGF NLRKREQTLLQVMETLLPK APIGIFPLKHLFPRGIHQDS

CPn_0772 872400 870469
uvtD-DNA Helicase
KLGLIMTCISELNEAGRKAVTAPLNPVLVLAGAGAGKTRVVTYRILHLINGGIAPREILA
VTFTNKAARELKERI/NGCASTNEEDVEPVCTFHSLGVFILRSINLLNEENNETIVDQS
DAERLIBHAL-GHBLERHLADFIDAHE TOAHU-ANDELEFICHUS UPV TOFFCHYOK STLEKELRSVKGPGEKIRLFLGSTDREEADFVAAEILOLHRVONIKLEDICIFYRTNSOS
RTFEDALLRRIPYEIIGGLSFYKÄKEIQDILAFLRIFISKSDIVAFDRTVNLPKRGIGS
TTIFALTOYAIAGGLPILKACQO/LDTKDVKLSKKQQEGLQEYLALFPQIEHANTISLR
DFIESVVRITOYLEILKEDADTYKDRKSNLEELYHKALESEQQNPKTHLELFLDDLALKG
SDDDLNLTADRVNLITLHNGKGLEFRVSFLVGLEEQLLPHANSLGGTYENIEEERRLCYV
GITRAQDLLYLTAAQVRSLWGTVRMKPSRFLKEIPKDYMIQVR

CPn_0773 8/2485 873195
ung-Uracil DNA Glycosylase
FMONATIDOLPVSWOEOLPLCWREQLKEEWSKPYMOQLLIFLKQEYKEHTVYPEENCVFS
ALRSTPFOVTVVIIGOPPYGKGQAHGLSFSVPEGGRLPPSLINIFRELKTDLGIENHK
GCLQSWANGGILLLVYVLTVRAGEPFSHAGKGWELFTDAIVTKLIQERTHIIFVLWGAAA
RKKCELLFNSKHQHAVLSSPHPSPLAAHRGFFGCSHFSKINYLLNKLNKPMINWKLP

CPn 0774 873183 873425

CFICE/TVA
CT506.1 hypothetical protein
LEAPHNEGIHS/CFOKTPRLTAKSVVSMEMLLTTQQLPSAEGMPSVANLEADFLRAEALL
AEMREIRGCLEOSLRTLVPSE

CPn_0775 874040 873414
yggV famfily
ERFMKIV/ASSHGYKIRETKTFLKRLGDFDIFSLSDFPDYKLPQEQGDSITANALTKGIH
AANHLG/WIADDTHLRVPALNGLPGPLSANFAGVGAYDKDHRKKLLDLMSSLESLVDRS
AYFECCVVLVSPNGEIFKTYGICEGYISHQEKGSSGFGYDPIFVKYDYKQTFAELSEDVK NOVSYRAKALOKLAPHLQSLFEKHLLTRD

0776 874180 875487

CPy_0776 874180 875487
CTp605 hypothetical protein
FifvLknfydclmffoflsftmkkifysfvllscifpyvgcaqvfvgldrifsegeytr
Figkkialishsaxinsrgopalsvfysgrkdctveilctlehgyvgatfretvgnops
Rypmlrsvslydykepykevaehcdvfyvdvgdigysfysfvtvlmotykaserykgcli
Vldrpnpmgcriudgelnmftsgslaipycvgmfrgelalffkktyapnavvvipmkg
wnrsmffdetglimmftspgmpdpdspffyaatgilgalsvasigvsttlffkvlgapm
dgekvadelnmkklpdvflpffyfspffkykmexcsgvllvlqdpkffypvetqctimg
Vlkalypkqveotlksieriparrssicnlfgddeflsishkeryivwplrrlckesres FHOLRSSCLLSEYAES

CPn_0777 875586 877178
groEL_2-heat shock protein-60
TSEDRVWWFKSQFEGLSALKRGVHALTKAVTPAFGPRGYNVVIKKGKAPIVLTKNGIRI
AKEIILQDAFESLGVKLAKEALLKVVEOTGDGSTTALVVIDALFTQGLKGIAAGLDPQEI
KAGILLSVEMVYQQLQRQAIELQSPKDVLHVAMVAANHDVTLGTVVATVISQADLKGVFS SKOSGISKTRGLGKRVKSGYLSPYFVTRPETMDVVWEEALVLILSHSLVSLSEELIRVLE
LISEQNTHPLVIIAEDFDQNVLRTLILNKLRNGLPVCAVKAPGSRELRQVVLEDLAILTG
ATLIGQESENCEIPVSLDVLGRVKQVMITKETFTFLEGGGDAEIIQARKQELCLAIARST
SESECQELEERLAIFIGSIPQVQITADTDTEQRERQFQLESALRATKAAMKGGIVPGGGV
AFLRAAHAIEVPANLSSGMTEGFETLLQAVKTPLKVLAQNCGRSSEEVIHTILSHENPRF
GYNGMTDTFEDLVDAGICDPLIVTTSSLKCAVSVSCLLLTSSFFISSRTKT

CPn_0778 877400 878092
tsa/ahpC-Thio-specific Antioxidant (TSA) Peroxidase
APVAgSDRVEGYEPGGQRRESSLVRNNKRVEEEVFMILDLVGKEAPDFVAQAVVNGETCT
VSLKDYLGKYVVLFFYPRESTYVEPTELHAFQDALGEFHTRGAEVICCSVDDIATHQQWL
ATKKKQGGIEGITYPLLSDEDKVISRSYHVLKPEEELSFRGYFLIDKGGIRHLVVNDLP LGRSIEEELRTLDALIFFETNGLVCPANWHEGERAMAPNEEGLQNYFGTID

CPn 0779 878502 878095

878502 878095
CT602 hypothetical protein
RFDLIFGMKFTVALFGEAEKGSYDTAYFCRSLVDLHNYLGDVSSPGITLAIKTLLSDYNV
VYFRVREEGYCVDSYFFGLHFLNYOTTLKNIIAIGLPGVGNQHIIEASRSLCQKHNSLLL
FFDHDLYDLLTFNQPF

CPn_0780 879241 878591
papQ/amiB-N-Acetylmuramoyl-L-Ala Amidase
HGNKIAVOSLRFMHAKLSFFILLSLLFSGIDCSFLHAAGRSPSLQOVLAEIEDISAKLAS
HEVEIVMLSERLDEQDSKCQKWTAAKPETLAQKIRELESLQKALAKTLAVLTTSVKDLQT
NLOSKLOEIQKDHRALAQDLRLVRRSLLALVDSSSPGAYADFSDPVPENIYIVREGDSLS
KIAKKYKLSVTELKKINKLDSDAIYAGQRLCLQPNKQ

CPn_0781 879851 879198
pal-Peptidoglycan-Associated Lipoprotein
ONCYRGRRKTYPLLGCFPSATOKENTMINISLWHLCTLLALLALPACSLSPNYGWEDSCN
TCHHTRRKRPSSFGFYPLYTEEDFPNPTFGEYDSKEEK/YKSSQVAAFRNITFATDSYT
IKGEENLAILTNLVHYMKKNPKATLYIEGHTDEPGAASYHLALGARRANAIKEHLRKQGI
SADRLSTISYGKEHPLNSGHNELAWQQNRRTEFKIHAR

379773 CPr. 0782 881077

CHI_0/82

SELOT 9/9/13

COLEMBROACCHAPIGE TRANSPORTER
CDIT/MEMOLEFOVFFFCFASLDYABELEY/VMSEHITLPTEVSCOTD/KDPKTOKYLSSL
TELFCKDIALIDKLA/PTAASKESSSPLATSLEHHVPQLS/Y/LLQSSKTPQYLCSFTTEQN
LSVDP/KTHHAADPVHVALTSTPGTSAGKTVFALSSLSYDOKLKQSELWTDYD/KNLAP LENDRYK THIMADITY TYALTYT POLSMAN TYRALDSOON DANDON HERF LYTTD ANNIAGE LENDRYS LISTTPRKWANGENFPYLYNGYMYGOPP LFLOSLEITPEGKKVLPLIK INOLMPTES PRKKLLAPVADTYGROPLE LOPESLTSOPMOR PERLLEITPOTONNEISPIRFEGIOLDF LS NKUBEREN THIMADITY EN ALTITIKKYENSGCEAMS POOKKLAPCISY (KGVRQ LC LYDL). COMPANYOUTE STEENES ESWALDERHLY FOAGNAEES ELY LIGEVYKKTNK LA IGVOEKRE POWGAFFOOP I KICTI.

RRIBBR 881109 8 hypotherical protein

EMMKAPIA LY LAUCHEST FOR NEW TOPICK KETÖBKYŁ ČEKPAL I ÓDKU NALIESTANDU.

AKT IRPSVATOPOKOAKCSPPOENVOKALOKP I PKV I P SPAPTVAKKTTATEKP WPSTAQLTMHSELKAT PRSTTKKNTQLGKTQLQTLGEVAQALGLHVDK1EKSET QEDELCELFRTHIALPSKGYVRIKLVLGPNGEIQECSFL SAADKOLLTORIOALPFO KFLEKYKVSKNISFHIKLVSNES

982359 881892 exbD-Biopolymer Transport Protein
prabsiftTeFYPO0IPDYKLMKYPFTEEIEEEPLVNLTPLIDIVFVILMAFIVAVPLIK CONTADADO POR JEVILO ENDO TAVA E VEADRO CIONERE OTROCENTA ELIMENA O THEFFORGETS ESTACHAND THE WAS REFIELD AND W

883039 CPn_0785 CPH_0785 883039 882296

exbB/tolQ-polysaccharide transporter

DHLYFETLSVNKDFYSMVHFSHNPIIQAYTEADFFGKSIFFCLLILSVCTWTVLHQKLAI

QKNFLKAGKSLKDFLIKNRHAPLSLDIHEELSPFADLYFTIKRGTLELLDKNRQSAPDRG

PILSSEDIOSLETLLGAIMPKYKALLHKNSFIPATTISLAPFLGLLGTVWGILVAFTHIS

SGSSGNSAIMEGLATALGTTIIGLFVAIPSLIAFNYLKAHSSELISEIEQTAYLLLNSIE VKYRNTNI.

CPn_0786 883137 885293
dsbD/xpra-Thio:disulfide Interchange Protein
NHGVILNKFKTYLOTALIAPFFSFPALSGSFSSIOAEEITOQVNHPGAELLSBGSYIPGL
QTFRLGIKITASKGSHIYWKNPGGIGSPLKISWGDPKGFVVEEEHWPTPKVFEEEGTTFF
GYEDSALIVADVRAPEGYTFGGEVELRAQVEWLACGDSCLPGNVDLKLTLPYEEKEPSLY
PDTHAAEFTKTLHAQPRVLENDHSVQVAQGKGMEIILNISKKINATKAMFVSEKADKLFAY
AETSYSGGTGTAWRLKVKNLSGVQKNEKLHGILLADHTGRPVESLTHSEVLGQTGSAV
AGLSQYITILIMAFLGGVLLNIMPCVLPLVTLKVVGLIKSAGEHRSSVIANGLWFTLGVV
CCFWGLAGVAFILKVLGHNIGWGFQLQEPMFVATLIIVFFLASIGLFFBGTMFANLG
GKLQSSEMKSSNNKAVGAFFNGILATLVTTPCTGPFLGSVLGLVMSLSFLQQLLIFTAIG
LGMASPYLVFSVFFKMLSVLPKPGGMMSTFKQLTGFMLLVTVTVMLVMIFGSETSTTSVV
LLGGIMLAGIGAMILGRWGTPVSPKKGVCVASLLFFAFLGGGISVSGLASTYFAEPQGTV
SVNEDSLWQPFSLEKLAQLRAQGRPVFVNFTAKWCLTCQMWKPVLYGDAVQKWFETHGIV
TLEADWTRKDPGITEELARLGRASVPSYVYYPGDNSAPVVLPEKITQNLLEDVVSRFVR

CPn_0787 885604 986401

yabD/ycfh-PHP superfamily (urease/pyrimidinase) hydrolase

TRROPVDLADAHVHLSDDAFEEDINSVLQRAQDSGVSLVWNVTTTEKELNRSFAYAERFP

KIRFCHVGGTPPQDVDQDIEEDYRNFHAAAHSKKLAAIGEVGLDVCFATEGSIARQKEVL

QRYLALSLECELPLUVHCRGAFNDFFRHLDQYYRNDPRSRPGMLHCFTGTLEEAQELISR

GWFISISGIVTFKNAQDLRDLVVELPLEHLLIETDAPFLAPVPYRGKKNEPAHVLHTINA VANVKGMFPQELAALAYKNVLRFLHG

CPTE0788 886521 887432
sdlc_Succinate Dehydrogenase
SLWSLRMSRHEICPEVSHKKGKYYSTFIFRCIHSLAGIAFTFICEHLFTNMLASSYFS
CGKGFVAMVNGFHKIPGLKIIEVAGLVLPFLCHAIIGIVYLFGGKSNCYSGDGSRPHLRY
AKMYSYTWORWTAWILLFGIAFHVVHLRFIRYPVHVDIHGTTTYAVDIOPSRYDVIVRGT
KGELTLNLPNTEASSIEVSRHDLGGADAALLSERNSYLLTPSAGTAFLYVVRDALGSLFI
ALÄMTILVIAAAFHGFNGLWTFCCRWGVVVSLRMQGVLRIVCYLAMIVVTFMGVSAVWNL
YSVA

CPH-0789 887436 889316

SdhW-Succinate Dehydrogenase

OMDENRKVIVVGGGLAGLSAAMQLANLGIIVELVSLTKVKRSHSVCAQGGINAALNLKPE

EEDSPYVHAYDTIKGGDFLAADPPVLEMCLAAPRIIKMLDNFGCPFNRGPSGNLDVRRFG

GTLYHRTVFGGASTGQQLWYTLDEOVGRREHAGRVIKRENNEFVALVTDHSGRACGIILM

NLFNNRLEILRGDAVIIATGGPGVIFKMSTNSTFCTGAANGRLFLOGMAYANPEFIQHH

TABBGRKKLRLISESVRGEGGRWVPGDSSKRIVFPDGSEPCGETGAPWYFLEDMFPAY

GNLVSRDVGARAILRVCLAGLGIDGMEAYLDVTHLPEKTRIKLEVVLDIYKKFTEDDFN

TVFHRIFPAVHYSMCGAWVDWPAADDPDRDSRFROMTNIPCGFNCGESDFOYHGANLGA

NSLLSCLFAGLVSGDEASRFIEAFGASQATSSDFDRALOGKEENARLLSASCKENIFVL

HEBTAK HWRNVTVKRNNRDLQETMBKLKEFRERLKNVSVLDSSFFANKSFHVROMGPW

LELALAITKGALLRNEFRGSHYKPEFPERDDEHWLKTTVAVYAPEEPEISYLPVDTRHVA

PTERDYTKSSTGKIELTNIPDNIRLPI

889279 890103 CYMEN OF THE PROPERTY OF THE P CPKKLPLTESISAVGREISKFSLRSLFSALFKKKK

893104 890111 CPn_0791 CFM_0791 893104 890111
CT590 hypothetical protein
TCLRSSRKIVVEDISDRNMYSCYSKGISHNYLLHPMSRLDIFVFDSLIANQDQNLLEEIF
CSEDTVLFKAYRTTALOSPLAKNLNIARKVANYILADMGEIDTVKLWEAIHHLSQCTYP
LGPHRNHEAQDREHLIKMLKALKENPKLKESIKTLF VPSYSTIQNLIRHTLALHPOTILS
TIHVRQAALTALFTYLRQDVGSCFATAPAILIHOFY PERFIKDLNDLISSGKLSRIVNOR
EIAVPINLSGCIGELFKPLRILDLYPDPLVKLSSSFGKKAFSAANLIETLGDSEAQIQQ
LLSHQYLMOKLQNHHETLTANDIIKSTLLHYYDLQESTVRAIFFKEGLFSKEQVAFSTQH
PRELSEIGNYYHYLHAYEEAKSAFIHDTONPLIKAMEYTLATLADASQPTISNHIRLALG PRELSELORVYHYLHAYEEAKSAFIHDTONDELKAWETTLATLADASOPTISNHIRLALG WKSEDPHSLVSLVTHFVEEEVENIRILVOOZEOTYHEARSQLEYIEGRMRNPLNNQDSQI LTMDHMFROGLINKALYEMDSAQEKAKKFIHLPEFLLSFYTKQIFLYFRSSYDAFIOEFA HLYANAPAGFRILFTHGRTHPNTWSFIYJINEFIRFLSFFTSTESELLGKHAVINLEKE TSRLVHNITAMLHTDVFQEALLTRILEYOLPVPPSILNHLDQLSQTPWVVVSGTVDTL LLDYFESSEPLTLTEKHPENPHELAAF/ADALKDLPTGIKSYLEEGSHSLLSSSPTHVFS I LAGSPLFREAWDDWYSYTWLRDWWKGHOPELDTILPQLSIYAFIENFCNKYALOHV VHOFHDFCSDHSLTLPELYKOSFFLSSLFTKDKTVALIYIRRLLYLWVREVPVVSEQQL FEVLDNYGSYLGISSRITYEKFRSYLEETIPKMTULSSADLRHIYKGLLMOSYOKIYTEE LTYLRITTAMRHINLAYPAFLLFADSNWPSIYFGFILIPGTTEIDLWKFNYAGLQGOPLD HIGELFATSHPWTLYANFIDYSWPPPRYRSRLPKEFF

893108 CPH_0792 894/\$5 893108
CT589 hypothetical pystein
RHHLINIKGIRIMKEPTKRYLEFFFLVIPIPLLLHIMVVGFFGFGAAKANLVQVLHTRA
THISTIEFERKLT HKLFLDKLANTLALKSYASPSAEPYAQAYNEMMALGHTDFSUCLIDP
FCCOVRTYHIGDDETRYLKOHPEMKKKISAAVGKAFLLTTICKPLLHYLILVEDVASWDS
TTTCSCLLVSHYPEMFGLQADLFGSEHTTKCNICLVMYYGEVLEGAQDRESSFVFGDLPNL
FGFQARSHSAIETEKAAGILGENLITVSINKKRYLGLVINKIPIQATYTLSLVPVSDLI
QGALKVPLHICFFYVKAFLIMMWIFSKINTKLNKPLGETFYMEAAWRONHNVRFEPQPY GYEFNELGHIFNCTLLLLLIA DIDYHGGEKLÖKEZÖTLÜĞÜLÜNDÜLÜNDLÜPDFFFFKVV TESSOHLRERQLSGHENWIT SOTLLGTISLLADYĞLDBYYLYALISARILELAYASSD VILQKISKOTNOSFOKTTEGIEAVVANTFIKYVEKDRÜLELLÜLÜSEGAFTNELQRGESFV RLPLETHOALQPODRLICLTYSEDILKYFSQLPIŞELLKDELIFLIYTENLIDSLTMHLINI ETEHSADGTLT LLSF3

996838 99491 rbsU-sigma regulatory family protein-PP2C phosphatase (RsbW

CONTROL OF THE STATE OF THE STA

697123 CPn_0794 898004

CPn_0795 898008 899195
NO FODUST NOMOLOG PRESENT IN GENEBARK/EMBL AS OF 11/7/98
GTLGGANSSATGVSSDGSVIVGQAQTADKSVHAFQYYNGEMKDLGTLGGTSSTAKTVSPD
GKVIMGRØGIADGSWHAFMCHTDFSSNIVLFDLDNTYKTLRENGRQLNSIFNLQNMMLQR
ASDHEFYEFGRSNIALGAGLYVNALQNLPSNLAAQYFGIAYKIRPKYRLGVFLDHNFSSH
VPNNFJVSHNRLWMGAFIGWQDSDALGSSVKVSFGYGKQKATITREOLENTEAGSGESHF
EGVANDIEGRYGKSLGGHVRVQPFLGLQFVHITRKEYTENAVQFPVHYDPIDYSTGVVYL
GIGSHIALVDSLHVGTRMGMEQNFAAHTDRFSGSIASIGNFVFEKLDVTHTRAFAEMRVN
YELPYLQSLNLILRVNQQPLQGVMGFSSDLRYALGF 898008 899195

CPn_0796 899280 901340
No robust homolog present in Genebank/EMBL as of 11/7/98
SELYSSYLOPCLNMSIVRNSALPLPCLSRSETFKKVRSHMKFMKVLTPWIYRKDLMVTAF
LLTAIPGSFAHTLVDIAGEPRHAAQATGVSGDGKIVIGHKVPDDPFAITVGFGYIDGHLQ
PLEAVRPQCSVYPNGITPDGTVIVGTNYAIGMGSVAVKWNGKVSELPMLPDTLDSVASV
SADGRVIGGRNININGASVAVKWEDDVITOLPSLPDANNACVNGISSGGSIIVGTMVDV
SMRNTAVQWIGGDLSVIGTLGGTTSVASAISTDGTVIVGGSENADSGTHAYAYKNGVMSD
IGTLGGFYSLAHAVSSDGSVIVGVSTNSEHRYHAFQYADGQMVDLGTLGGPESYAGVSG
DGKVIVGRAQVPSGDMHAFLCPFGAPSFAPHGGSTVVTSQNPRGMVDINATYSSLKNSQ
OQLCRLLIGHSAKVESVSSGAPSFTSVKGAISKQSPAVQNDVQKGTFLSYRSQVHGNVQN
QQLLTGAFMDWKLASAPKCGFKVALHYGSQDALVERAALPYTPGGGLGSSVLSGFGGQVG
QQLTGAFTDWKLASAPKCGFKVALHYGSQDALVERAALPYTPGGGGSSVLSGFGGQVG
RYDFNIGETVVLQPFMGIQVLHLSREGYSEKNVRFPVSYDSVAYSAATSFMGAHVFASLS
PKMSTAATLGVERDLNSHIDEFKGSVSAMGNFVLENSTVSVLRPFASLAMYYDVRQQQLV
TLSVWMNQQPLTGTLSLVSQSSYNLSF

CPn_0797 901552 902694

No robust homolog present in Genebank/EMBL as of 11/7/98

VLILTWINVLTKLGLMMSKKIKVLGHLTLCTLFRGVLCAAALSNIGYASTSQESPYQKSI
EDWKGYFTPTOLELLSKEGMSEAHAVSGNGSRIVGASGAGGGGSTAVIWESHLIKHLGTLG
GEASSAEGISKDGEVVVGWSDTREGYTHAFVFDGRDMKDLGTLGATYSVARGVSGDGSII
VGVSATARGEDYGWVGVKWEKGKIKQLKLLPQGLWSEANAISEDGTVIVGRGEISRNHI
VAVKWNKNAVYSLGTLGGSVASAEAISANGKVIVGWSTTNNGETHAFMHKDETHHDLGTL
GGGFSVATGVSADGRAIVGFSAVKTGEIHAFYYAEGEMEDLTTLGGEARVFDISSEGND
LIGSLKTHAFAGEAERAYLFHHKK IIGSIKTDAGAERAYLFHIHK

CPn_0798 902810 903856
No robust homolog present in Genebank/EMBL as of 11/7/98
vvFEIIFVVRVPMKKTCCQNYRSIGVVFSVVLFVLTTQTLFAGHFIDIGTSGLYSWARGV
SGORRVVNCYEGGNAFKYYDGEKFLLEGLVPRSEALVFKASYDGSVIIGISDQDPSCRAV
KWYNGALVDLGIFSEGMQSFAEGVSSDGKTIVGCLYSDDTETNFAVKWDETGMVVLPNLP
EDRHSCAWDASEDGSVIVQDAMGSEEIAKAVYWKDGEHLLSNI FGAKRSSAHAVSKDGS
FIVGEFISEENEVHAFVYHNGVIKDIGTLCGDYSVATGVSROGKVIVCHSTRTDGEYRAF KYVDGRMIDLGTLGGSASFAFGVSDDGKTIVGKFETELGECHAFIYLDD

905001 903940 CPn_0799 905001 903940
No robust homolog present in Genebank/EMBL as of 11/7/98
KREENMAAIKOILRSMLSQSSLWMVLFSLYSLSGYCYVITDKPEDDFHSSSAVKMDHWGK
TTLSRLSNKKASAKAVSGTGATTVGFIKDTWSRTYAVRWNYWCTKELPTSSWVKKSKATG
ISSDGSIIAGIVENELSQSFAVTWKNDEMYLLPSTWAVQSKAYGISSDGSVIVGSAKDAW
SRTFAVKWTCHEAQULPVCWAVKSVANGSVANGSIIVGSVDDASGILVAVKWEGNTITHL
GTLDGYSAIAKAVSNNGKVIVGRSETYYGEVHAFCHKNG'MSDLGTLGGSYSAAKGVSAT
GKVIVGMSTTANGKLHAFKYVGGRMIDLGEYSWKEACAHAVSIDGEIIVGVQSE

906550 905249 eno-Enolase eno-Enolase
RKEIKIMFEAVIADIQAREILDSRGYPTLHVK/TTSTGG/GERVPSGASTGKKEALEFR
DTDDPRYQGKGVLQAVKNYKEILFELVKGGSV/EGSLIGSLMMDSDGSPNKETLGANAIL
GVSLATAHAAAATLRRPLYRYLGGCFACSLPCPMMNLIN/ZMHADNGLEFQEFMIR PIGA
SSIKEAVNMGADVFHTLKKLLHERGLSTGVGDEGGFAPHILASNEEALELLLLAIEKAGFT
PGKDIGLALDCAASSEYNVKTGTYDGTRYEEGIALLGHLCDRYPIDGIEGGLAEEDYDGW
ALLTEYLDEKVGIVGDLFVTNPELILEGISNGLANGVLIKPNQIGTITTETVYAIKLAQM
AGYTTIIGHRGGTTTYTIADLAVAFNAGQIKTGGLSGERVAKYNRIMEIEEELGSEAI PTOUNTERVEDUEE

908709 906727 CPI_9801 908709 906727

over eximplement and subunit b
ITEMMEDIATE AND Subunit b
ITEMMEDIATE AND SUBUNIT B
ITEMMEDIATE AND SUBUNIT B
ITEMMEDIATE AND SUBURIT B
ITEMMEDIATE AND SUBURIT B
IDKLEDSAME I LERROMITA SUBURIT SUBURIT B
IDKLEDSAME I LERROMITA SUBURIT SUBURIT SUBURIT B
INTERNATION OF SUBURIT SUBU

SGH I'VQQI IRPTG IPDP GFLSELE IPAAYLHSGI VEYOFRUPSAFONRPLTYEEAOKYFRKZIYVSATPGD1 WETTRATTOYDOLLEETRLPLOCKHEKTLYTTITKK GFLSELTRAYLINGT ETAERTOTUTDLRGGVIDVLIGONLLREGLDLPEVSLVA, JADKSGFLRSTSSLIGFG RAARNINGKVIFYADOKTRSIEETLRETERROIQLDYNKENNIVPKPIIKAIFANPILQ TSKDSESPKESORPLSKEDLEEQIKKYEALMGRAAKEFRFNEAAKYRDAMGACKEQLLYL

202761 208703

CPI_0802 909761 908769

THE TYPE SHAPE SHA KERRSEFLSKPLALONVLEDGTHKMREVAKVTMEEVHDKFGFSHKWRSLLK

909752 910306

CFI_0003 970206 CT584 hypothetical protein FMAKKKTLELEENVFLLLEGKIKRIFATPIGYTTFREFQNVVFNCANGQQEIANFFFEM LINGKLTQELAPQOKQAAHSLIAEFMMPIRVAKDIHERGEFINFITSDMLTQQERCIFLN RLARVDGQEFLLMTDVQNTCHLIRKLLARLLEAQKNPVGEKNLQEIQEEITSLKNHFDEL

CPn_0804 911074 910310
gp6D-CHLTR Plasmid Paralog
EIFSSMGNLKTLLESRFKKNTPTKMEALARKRMEGDPSPLAVRLSNPTLSSKEKEQLRHL
LQHYNFREQIEEPDLTQLCTLSAEVKQIHHQSVLLHGERITKVRDLLKSYREGAFSSWLL
LTYONROTPYNFLVYYELFTLLPEPLKIEMEKMPRQAVYTLASRQGPQEKKEEIIRNYRG
ERKSELLDRIRKEFPLVETDCRKTSPVKQALAMLTKGSQILTKCTSLSSDEQIILEKLIK KLEKVKSNLFPDTKV

CPn 0805 911846 911067

FILE OF THE PROPERTY OF THE PR TILGVALSFWNCRGKNNSAFAELIHKTFPGKLLNTKIRRDITVSEAAIHGKPVFATSPSA RASEDYFNLTKELLILLRDI

CPn_0806 913816 911867

LHTS-Threonyl LRNA Synchetase
NÄNNESPPNMEAMNKMIQVTCDOKNYEVLEGTTAAELAKOLKNSHQFIGVLINERPROLS
THLHEGDTLVFLTSEDPEGREIFLHTSAHLLAQAVLRLWPDAIPTIGPVIDHGFYYDFAN
LSISSDFPLIEDTVKQIVDEKLAISRFTYGDKQQALAQFPQNPFKTELIRELPENEEIS
AY\$GCEFFDLCRGPHLPSTAHVKAFKVLRTSAAYWRGDPSRESLVRIYGTSFPTSKELRA
HLEQIEEAKKRDHRVLGAKLDLFSQQESSPGMPFFHPRGMIVWDALIRYWKQLHTAASYKE
ELEPPCLNNROLWEVSGHWDNYKAMMYTLQIDDEDYAIRPMNCPGCMLYYKTRLHSYKEF
PLRVAEVGHVHRQEASGALSGLMRVRAFHQDDAHVFLTPEDVEEFTLNILQLVSTLYGTF
GLEYHLELSTRPEKDTIGDDSLWELATDALNRALVGSTPFIVRPGEGAFYGPKIDIHVW
DAIQRTWQCGTIQLDMFLPERFELEYTTAQGTKSVPVMLHRALFGSIERFLGILIENFKC
RFPEMLSPEOVRIITVADRHIPRAKELEEAWKRLGLVVTLDDSSESVSKKIRNAQNMQVN
WITTLGDHEINENVLAVRTRDNRVINDVSVERFLNTILEEKNSLSLTALL

CETTO807 913950 914879
CTT50 hypothetical protein
TLOTGLHMSLFLVFLTAFIWSSSFALSKLVMNASAPIFATGARMVIAGAILALAAWFRGS
FVGISKKIFLYIVLLALTGFYLTNIFEFIGLOSLSSSKTCFIYGLSPLMSALFSVIQUE
KVFEKKVIGLISLGLVSYICYLTFGGGGDDSQPHTWOIGLPELLILGAASLASFGWTLKO
IEKQSTLSVTAINAYAMLIAGMLSIMHSAVVEPWRPLPVQDISQFLYATLALVVISMIIC
YNLTAKLLRKYSSTFLSFCNLVMPLYSGFYGWILLGEKGVSLGLVLAVAFMVAGCELIYH EEFROGYIVS

CPALO808 916398 914956
CT579 hypothetical protein
LKKEPSWALKSLKKRMPOSAEPSLAHIKPIIFKGACIAMTSGVSGSSSODPTLAAQLAOSS
CKAGNAQSGHDTKNVTKQGAQAEVAAGGFEDLIQDASAQSTGKKEATSSTTKSSKGEKSE
KSGKSKSSTSVASASETATAQAVQOPKGLRQNNYDSPSLPTPEAQTINGIVLKKGMGTLA
LLGTWNTLMANNAGESWKASFQSQNQAIRSQVESAPAIGEAIRRQANHQASATEAQAKQS
LISGIVNIVGEPTVSVGAGIFSAAKGATSALKSASFAKETGASAAGGAAJKALTSASSVQ
QTMASTAKAATTAASSAGSAATKAAANLTDDMAAAASKMASDGASKASGGLFGEVLNKPN
WSEKVSRGMNVVKTQGARVASFAGRALSSSMGMSQLMHGLTAAVEGLAAQGTGIEVAHHQ
RLAGQAEAQAEVLKQMSSVYGQQAGQAGQLQEQAMQSFNTALQTLQNIADSQTQTTSAIF
N

CPn_0809 917794 916307
CT578 hypothetical protein
DTNMSISSSSGPDNOKNIMSQVLTSTPQGVPQODKLSGNETKØIQQTRQKNTEMESDAT
LAGASGKDKTSSTTKTETAPÇQGVAAGKESSESQKAGADTGYSGAATTASNTARTYARG
TSIEEASKSMESTLESLOSLSAAQMKEVEAVVVAALSGKSSGSAKLETPELPRKGVTPAS
EVIEIGLALAKAIQTLGEATKSALSNYASTOAQADOTNKLELEKQAIKIDKEREEYQEMK
AAEOKSKOLECTMDTVITVMIAVSVAITVISIVAAIFTCAGLAGLAGLAGAAVGAAAAGGA
AGAAAATTVATQITVQAVVQAVKQAVITAVRQAITAAIKAAVKSGIKAFIKTLUKALAKA
ISKGISKVFAKGTQMIAKNFEKLSKVISSLTSKWVTVYGVVVVAAPALCKGIMQMQLSEM
QONVAQFOKEYOKLQAAADMISMFTQFWQQASKIASKGTGESNEMTQKATKLGAQILKAY
AAISGAIAGAHKTNNF

918193

CPD_UBIO 918193 917825 CT577 hypothetical protein GEIMMKKMKKTKKAVQSKAAPVKRVPEESQEAATQQLELAVSDLYKELPLAQTFASLTDK NQ THSTTAALSGTLESLHH.EELTQGLFFSAQERANFAKELSSVVHGLKNLTTVVNKQMVK GAE

918933 218/208

Clac_9812

920862

MULL-DNA MISMALCH R
GILICWLCNLTKAPMSTRRI
DPLT NO LAAGEVIENSVOVVKELIEISLDAGADET
EIETLOGCOGALIIEDNOCGFRAEDIF VALORHATSKIREFSDIFSLNSFSFRGEALPSI
ASISKMELOSSIECDEGVRTVIHGGDIVSCEPCAROLOTTVIVNSLFYNVPVRRGFOKSM
OSDRIGIRKLIENRILSTANIGWSWIJECHHEIQIAKQQSFQERVAYVMODHFWODALTI
DKEANGVRIVOVLOSPSFHRFTROCOKIFINDRPIESLFISKKVGDAYALLLPLHRYPVF
VLKLYLPSSWCDFMVHPOKIEARILKEELVGDIKEAIVETLACPPSILGRTHQEIEESD
SVPLFMFRMLETSDVQEEESVEFDONLFAYSSEDVSLEKOFYTSGGPKSOMDWIYSSDV
RETTSJRVZLARDLSTANILIETVAPPHLEBLAGS, ROLBENT, LAAGAGETE, INSEALT
ALMERITIONAFSKUDVERDVSMIKELBEVOKEREGEDIAAARARGETE INSEALT
ALMERITIONAFSKUDVERDVSMIKELBEVOKEREGEDIAARRILETRILLDSDFMG mut L-DNA Mismatch R ALMKETLTQATFSKHQHVFDV9WLKLLWSVGKPEKGFDJARIRRLILDSDFMEG

CPn_0813 92/843 921934
pepP-Aminopeptidase/P
TLILWKDNHAMDRILRAGRALSEHNLDAILVEKSEDLAYFLHDEAIAGILLIGQGEVMF
FVYRMDKDLYSHIGRVPLYTELTODVVADLSLYVOKGRYCKIGFDSASTVYHKFAQRGVLP
CLWEPLECFTEKIRSIKSEESIRRMGEAALGSAGYDYVLTLLREGITEKEVVRQLRAFW
AEAGAEGPSFPPILAFEHSAFPHSIPTDRPLKKGDIVLIGUVLLNGVSDMTRHTALG
TPHPKLLESYPVVVEACKRAMALCKEGVLWGDIDAEAVRVLREHHLDTYFIHGIGHGVGR
HIHEYPCSPRGSQVYLESGMTITVEPGVYFFGIGGIRIEDTLCIDKNKNFSLTARPVISE

CPn_0814 921996 923357
CT814.1 hypothetical protein
FFLFFKLSYNFIFNLPLTMYQLLSIGYSFVSFIALLWMLCYSPNYVTDLYRISLSAEESL
GGIRAFPQAESLLGGACALNFPDLEERLPDLRKELLFLGSNDRPDACGGKFSLQLASSKE
CYIAALKERVILNYTNSSRGFVYSFSFRGVPTELMIECFSVSVDGRVEVKVRLQGLHKEL
ISKPRDCETLFLNPPANKLDCWEIAGFRVDASFPVKQKIRRIGVDKFLLMHGGAEYADKA
TKERVDFYSSDEENYSRVLAVGDVLLWGDCNCVTCGFFGASSRAPLFEVKRIDDKVMIA
DLWNVGCTQROTISLVKGVPSPIEINEVIRRIEFTGMRSWSKPIVLVGQGRLILSFDDMV
LRTAKGWEKLSRADQIGDYVTGKVTGPLLVFEKLEKDLRGFVLRGHMFNAQRTLVETISL
PLKCGFFEDVAVSORVSSNTRSAAAHPGATNRGCG PLKOGFEPAVASOEVSSNTRSAAAHPGATNRGGS

CPn_0815 923361 925622 gspD/pilQ-Gen. Secretion Protein D mvFrnSLLHLVALSGHLCCSSGVALTIAERWASLEHSGRGADDYEGMASFNANMREYSL QVSKLYEBERKLRASGTEDEALWKOLIRRIGEVRGYLREIEELWAAEIREKGGNLEDVAL WNHPETTIYNIVTDYGTEDSIYLIPQEIGAIKIATLSKFVVPKESFEDCLTQILSRLGIG WHEPTTIYNLVTDYGTEDSIYLIPQEIGAIKIATLSKFVVPKESFEDCLTQILSRLGIG
/RQVNSWIKELYMMRKECCSVAGVFSSRKDLEALPETAYIGFVLNSNVDAHTNOHVLKKF
INPETTHYDDYIAGRWIFGSAGEVGELLKIYNFVQSESIRQEYRVIPLTKIDRGEMISIL
NAAFREDLTKDVSEESLGLRVVPLQYQGRSLFLSGTAALVQOALTLIRELEGGIENFTDK
TVFWNVKHSDPQELAALLSQVHDVFSGENKASVGAADGCGSQLNASIQIDTTVSSSAKD
GSVKYGNFIADSKTGTLIMVVEKEVLPRIQHLLKKLDVPKMYNIEVLLFERKLAHBOKS
GLNLLRLGEEVCKKGCSPSVSWAGGTGILEFLFKGSTGSSIVPGYDLAYOFLMAQEDVRI
NASPSVYTMNQTPARIAVVDEMSIAVSDKDKAQVNRAQYGIMIKMLPVINVGEEDGKSY
ITLETDITFDTTGKNHDDRPDVTRNITNKVRIADGETVIIGGLRCKQMSDSHDGIPFLG
DIRGIGKLFGMSSTSDSLTEMFVFITPKILENPVEQQERKEEALLSSRPGEREEYYQALA
ASSEAARRAARKKLEMFPRSGYGISGVERDGFYUG ASEAAARAAHKKLEMFPASGVSLSQVERQEYDGC

925600 927102

gspE-Gen. Secretion Protein E RGKNTMAASILSQELLDILPYTFLKKHCLLPIEESSEAITIAHATATSVIAODEVKLLIK RGKNTMAASILSQELLDILPYTFLKKHCLLPIEESSEAITIAHATATSVIAQDEVKLLIK KPVRFVLKEESEILQRLQQLYSNREGNVSDMLLTMKEEDGTTISEEEDLLETTDTIPVR LLWILKEAIEERASDIHFEPCEDSMRIRYRIDGVLHDRHSPPSHLRSALTTRLKVLAKM DIAEHRLPQDGRIKIHIGGQEVDMRVSTVPVIYGERVVLRILDKRNVILDIAGLHMPKGT EILFKDTITAPEGILLVTGPTGSKTTTLYSVLQELKGPLTNIMTIEDPPEYKLPGIAQI AVKPKIGLTFARGLRHLRQDPDILMVGEIROGETAEIAIQAALTGHLVVSTLHTNDAI AIPRLLDMGIESYLLSATLVGVVAQRLVRTICPYCKVAYTPENQEKSFLASLGKDTEMPL YRCGCVHCFRSGYKGRGGITFELPNTLFRSEVASNRPYHILRETAEQNGFLPILEHGI ALAYSGETTLAFURRYRECD ALAVSGETTLAEVLRVTKRCD

CPn_0817 927106 928287
gspF-Gen. Secretion Protein F
GGRMPRYRYTYLDPKERKKRGYLEALHIQEAREKLAQENIQVLDIREVALRRMSIKSTEL
IVFTKQLLLLLRSGLPLYESLVSLEDQVHEORMGLLLTSFMETLRSGGSLSQAMAAHPNI
FDHFYCSGVAAGESVGNLEGCLQNIIVVLEERAQITKKMVGALSYPCVLLVFSFAVMLFF
LLGVIPSLKETFENMEVKGLTKIVFGVSDCLSAYRYLFLGFASALITVGILMRRIFWKK
LLEKLLFALPGTKKFVVKVAVNRFCSVASALLKGGGTLIEGLDLGCDAIPYDRLKTDMI
IVQAVIGGGSLSQELAQRSWYRKLAIGMIALGEESGDLADVLGYVAHIYNEDTQKTLASI
TSCHCODULI ISLGCI LGUJMIA IL IDITSNITGUT TSWCQPVILIFLGGLIGVIMLAILIPLTSNIQTL

CPn_0818 928158 928682
predicted OMP [leader (16) peptide]
GYTKNVGFDNVVVSTRDSDFSwWPDRCDHVGNIDPTHKQYPNIIKCVLRGVGMKRQKRKQ
SITLIEMMVVITLIGIIGGALAFNNRGSIHKGKVFQSEQNCAKVYDILMMEYATGGSSLK
EIIAHKETVVEEASWCKEGRKLLKDAWGEDLIVQLNDKGDDLVIFSKRVQSSNKK

CPn_0819 929117 928956 CT568 hypothetical protein ASLYGYCLFLIWEKFHNNIGKANFHLKIITTDFLTDIYIVTIRDPIAYPLTGIC

929042 929659

CPH_0820 929042 929659
CT567 hypothetical protein
DESLPCRCCCOTFFRSETSSIRTEMPMCNSIAMKKOKRGFVLMELLMSFTLIALLLGTLG
FWYRKIYTVOKOKERIYNFYIEESRAYKOLRTLFSMSLSSSYEEPGSLFSLIFDRGVYRD
PKLAGAVRASLHHOTKOORLELRICNIKDOSYFETORLLSHVTHVVLSFQRNPDPEKLPE TIALTITREPKAYPPRTLTYQFAVGK

CFn_0821 929637 930668

CPI_0821 929637 930668
CT566 hypotherical protein
HTMLELGINKEMOPP FETILICLIFOLVOLVAFDAANARKRCACAOTTERGENEEGIKRSACA
ELEVGEKORHAJA LERIOKUNGUVAFDAANARKRCACAOTTERGENEEGIKRSACA
ELEVGEKORHAJA LERIOKUNGUVAFDAANAL LONGOELEEGAAL
LLSEFJECYSIDTASWYA FETILIERRAYVOTTENI/PROSEYA FANAL LONGOELEEGAAL
LNSTJEEGAAE FEYRMLKISONOGOLLIFHIJYEERKOLGICKULLI FRODELLEAU
DHEPAYRETSILIRISTIMEAVIKROEHA FOERKOAAALELEKTRTDERLELRIKMOLLLSRY
DILIELLIKKMEDYTLISTA DYLELVDENTKA LORGREDIKKILLI

CEN_0822 43069710000 0.41229

CHRACESTRATIC CONTROL CONTROL OF THE

CPn_0823 932424

CPH_0823 93244 931501
yscT/apar-yopt Tranlocation T
FYALOVRFSKT: INCAKELMGISLPELFSNIGSAYLDY IFOHPPAYVMSVFLLLLARLLP
IFAVAPFLCAKLFPSPIKIGISLSMLAIIFPKVLADTOITNYMDNILFYVLLVKEMIIGI
VIGFVLAFPFYAAOSAGSITMOOGIOGLEGATSLISIEOTSPHGILVHYFVIIFMLVG
CHRIVISLLLOTLEVIPIMSFPPAEMMSLSAPIMTMIKMCOLCLVMTIQLSAPAALAML MODERN OF THE MAR OVEN EYEL SALE ARMODER OT LAWF LIKO LEVET LAWFK HAVE IMD

CPn_0824 932677 932378
yscS/fliQ-YopS/fliQ Translocation Protein
IRTRAVLAFFATSFKSVLFEYSYQSLLLILIVSAPPIILASIVGIMVAIFQAATQIQEQT
FAFAVKLVVIFGTLMISGGWLSNMILRFAGQIFQNFYKWK

932677 933618 CPH_U825

93518

932617

93518

932617

93518

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

932617

CPn_0826 934382 933612
yscl-Yop Translocation L
HDNRRSGVFSSEVNOPORYYAIVMKFFSLIFKDDDVSPNKKVLSPEAFSAFLDAKELLE
KTKADSEAYVAETEOKCAQIRQEAKDOGFKEGSESWSKQIAFLEEETKNLRIRVREALVP
LAIASVRNIIGKELELHPETIVSI ISQALKELTQNKHIIISVMPKDLPLVEKSRPELKNI
VEYADSLILTAKPDVTPGGCIIETEAGIINAQLDVQLDALEKAFSTILKAKNPVDEPSET
SSSTDSSSLSNDQDKKE

935273 934434 CT550 hypothetical protein GCLVTANTFGTLDILMKHSKEDDLSRFLPKNLLVESPHPEEIPLKSLSPTMSWLPTIHPS WITIANKEF PPEICOGLIAMLEPEDVOEILPLLEGOISIAPHRCAPFGAFYLLDMLSKKIR PCGITEEIFLPASSANAILYYTGPVKIALINCLGLYSIAKELKHILDKVVIERVKNALSP TEKLFLTYCQSHPMKHLETTNFLSSWTTDAELRQFVHKQGLEFLGKALTKENASFLWYFL RRLDVGRAYIVEQTLKTWYDHPYVDYFKSRLEQCMKVLVK

CPL 19828 936292 935267
yscJ-Yop Translocation J
IKEYAWIMVRRSISFCLFFLMTLLCCTSCNSRSLIVHGLPGREANEIVVLLVSKGVAAQK LPQAAAATAGAATEGMADIAVESAUTEALAILNOAGLPRIKEANEIVVLLVSKOVAOK LPQAAAATAGAATEGMADIAVESAUTEALAILNOAGLPRIKGTSLLDLFAKQGLVPSEL CEKERYQEGLSEGMASTIRKMOGVVDASVOJISFTTENEDNLPLTASVYIKHRGVLDNINS IMVSKIKRLIASAVPGLVPENVSVVSDRAAYSDITINGPWGLTEEIDVVSWGIILAKAS LTKFRLIFYVLILILEVISCGLLWVIWKTHTLIMTMGGTKGFFNPTPYTKNALEAKKATG AAABKEKKEDADSQGESKNAETSDKDSSDKDAPEGSNEIEGA

CPb≟0829 936729 937298 NO TODUST homolog present in Genebank/EMBL as of 11/7/98

KYĪĆPVPTLAKSFYINIRDSRFYSWLCFIMKETYYRDFLHENYLKNKKSMFHKIYKTAFE
FFLÄNAKWPLVPAGYRRYSKUFPVLSPLUDLVILLFPWVTKDSRYSPCSMTFTCICRS/VE
CIĘWYSTLFGIGRFCAVWCVEGFSGSTFDKIYHTIVAVLGILGIGILTFILRIIFS/TML **PVWFLFKCYS**

CPD_0830 937339 937959
NoFTObust homolog present in Genebank/EMBL as of 11/7/98
DSCSFLLPCFEVEAQTFPQVFSKVVVYKYKSSRILLIALLYNITLVLGLIFTHKKYLGQK
GRVTLKIYQNEEEFFRATERFPSIGAGYHKYNKNSVLFPFEDLMLVCPSYPKDFPLSAF
KVTTKLIYMSVLESIEVVGAFFFSIGRIFAMWCIEDFPGSIFSRIYHTIVSVLGILGLGI
IMFILRIIFTLLTLPFWLISCLKSSAA

CPI_0831 938249 938434
No LEGBUST homolog present in Genebank/EMBL as of 11/7/98
NKREANVLIRKSESEGAFFEATONYPTIQOGYOLVRIREHNLSVRAHFDLSLSLDASVHP
A4 11

939750 .938827 lipA-Lipoate Synthetase 11pA-Lipoate Synthetase
WMKCRPTLMTDOPRVRKKLPERPFWMLQRPLPQGSAFHATTATIKRSGMPTVCEEALCPN
RAECWSRKTATYLALGDVCTRSCGFCNIGHSKTPPALDPTEPERIALSAKELGLKHVVIT
MVARDDLEDGGAGGLVDIIQKLREELPOATTEVLASDFOGNVSALHTLLDSGITIYNHNV
ETVARLSPLVRHKATYARSMFMLEGGANYLPDDLKIKSGIMYGLGEMEGEVXOTLODLASI
GVBIVTIGQYLRPSRKHLQVKSYVTPETFDYYRRVGEMGLFVYAGPFVRSSFNADMILA SVQDKASA

CPn_0833 941171 939747

lpda-Lipoamide Dehydrogenase
RGVLFEILITVSENWTQEFDCVVIGAGPSGYVAITAAQSKLRTALIEEDQAGGTCLNRG
CIPSKALIAGANVVSHIKHAEQFGIHVDGYT7DYPAMARKKNTVVQGIRGGLEGLIRSNK
ITVLKGTGSLVSSTEVKVIGQDTTI IKANH YILATGSEPRFPRGVPFSSRILSSTGILEL
EVLPKKLAIIGGGVIGCEFASLFHTLGVEYTVIEALDHILAVNNKEVSQTVTNKFTKQGI
RILIKASISAIEESONQVRITVNDOVEEBGYVLVAIGROFNTASIGLDNAGVIRDDRGVI
PVDETMRTNVPNIYAIGDITGKWLLAHVASHGGVIAAKNISGHHEVMDYSAIPSVIETHE
EIAMVGLSLGEAEQQNLPAKLTKFPFKAIGKAVALGAGDGFAAIVSHEITOOILGAYVIG
PHASSLIGEMTLAIRNELTLFCIYETYHAHPTLSEVWAEGALLATNHPLHFPPKS 941171 939747

942014 hypothetical proty 'n KELIMPEANER HALL PROBERT KELIMPEANEREMORTOMKLEGSJYMIVPDOEPKSAFLJOPPVAGXSESSCHLOPPEGASK EGAEDLFAVSSEDMEAVLIDOMPTOETHKOVI PENTYLLOCMPLAALFLOIGLLAFAFLIL LESPIDGILVUTWEKNRAYEYYI IGAAVAYRGYRKLEL

45698 942045 CONTO STATE CONTROLL OF THE CONTROL OF T

DRENHPKTRIGSVEYVAKTHEMITGPKAIALPIYA
IPLLADKEKOOLLSLLCYOS
RYDIRLLADASESESAYLVTPODLONGSLIYPNYC
YSPTKGLMOVVGMLSPKOAFIVKSEOVEDFLJÆRGHLIQEFGFOTF INERPEGKLTYNYT
EOGVLLFHYDVODPSSTEIRFGTWTYYTNOFFLEKKNDLPIQDGLIVEPODIPAFIVKN
DAALRRLPNFFSSPPNLKDLLIEVHROSROKCLDLKPILVGLGESRCWLFGVFLYREDIG
FSLIPTPLQGLCFLPRVIPPENVPQFLTOYAQHERILFPNPQTRPPESYELVIQSIHRPH
PASPEHLOLELKTNLGSVPIGIALOGLISKHTFLFTOAGFLDLKONLFQFLKOFLSTOKC
VIAENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
VIAENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
UNENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
UNENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
VIAENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
VIAENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
VIAENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
VIAENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSSDKO
VIAENTVIANITOVEKLDALAPLSTYDDF IANPEDLOFFSOLKAACLPPIPONLFSKARTAV
VERDETHMAKKISSOLKH KILCKIDAMKALILTVIPPIPONLLESTYTTLENDALEFKOLDLILLINNIN PSALF THE THE PROPERTY OF THE PROPER

CPn_0836 946960 945722
brno-Amino Acid Branched) Transport
KMKKNASKKTNDKKSLSIWSIGGSIFAMFFGAGNIVFPLALGYHYNAHPWSAYFGMMLTA
VCVPLLGLVSMLFYSGDYOKFFFSIGRIPGMIFITAIILLIGPPGGIPRAIAVSHATLIS
LSEHKSAFIPSLPJFSAICCVLIYIFSCKLSRLIGWLGSVFFPIMLVTLLWVIIRSFMIP
THPWVQEFIPNAHQAWLAGFIEGFNIMDLLAAFFFCSIVLISLRQLVAEEKHFTEEEIPL
SFQGISKNRKSLALGFILAAILLGMITYLGFVLSAARHAGLLVNYSKGHILGRISAILL
SPGSISKNRKSLALGFILAAILGMITYLGFVLSAARHAGLLVNYSKGHILGRISAILL
SPGSISKNRSTALGFILAAILGMITYLGFVLSAARHAGLLVNYSKGHILGRISAILL
STGSILLAGVSVPIACLITEIALVGIVADFLARVVSFKKLNYASAVICTLIPTYLISILNFE
TISHLLLPLIQLSYPALIVLACGNIAYKLWNFRYSPVLFYLTLSLTIVLKLVN

CPn_0837 947145 nth-Enginuclease III LTMKQF/LRTLNALFPNPKPSLEGWSSPFQLLIAILLSGNSTDKAVNSVTPQLFAKAPDA OSILDLPPGKLYQLIAPCGLOERKSAYIYOLSQILVRDFHGEPPNDMALLTOLPGVGRKT ASVFLGIAYGKPTFPVDTHILRLAQRWKISEKKSPSAAEKDLARFFGHENTPKLHLQLIY YARQYCPALHHKIDNCPICSYLAKEANSTRT

CP1_0838 949196 947781 yriloso yndf-Thiophene/Furan Oxidation Protein ISLNIYPNSFHLFNLKLGILSESSFNFSIFMLKHDTIAAIATPPGEGSIAVVRLSGPQAI /ISLNTYPNSFHLFNLKLGILSESSFNFSIFMLKHDTIAAIATPPGEGSIAVVRLSGPQAI
VIADRIFSGSVASFASHTIHLGQVIFEETLIDQALLLLMRSPRSFTGEDVVEFQCHOGFF
ACSGILDALIALGARPALPGEFSQRAFLNGKIDLVQAEAIQNLIVAENIDAFRIAQTHFQ
GNFSKKIQEIHTLIIEALAFLEVLADFPEEGPDLLVPQEKIQNALHIVEDFISSFDEGQ
RLAQGTSLILAGKPNVGKSSLLNALLQKNRAIVTHIPGTTRDILEEGWLLQGKRIRLLDT
AGQRTTDNDIEKEGIERALSAMEEADGILWVIDATQPLEDLPRILFFKPSFLLWNKALD
PPPFLDTSLPQFAISAKTGEGLTQVKQALIQWMQKQEAGKTSKVFLVSSRHHMLQEVAR
CLKEAQQNLYLQPPEIIALELREALHSIGMLSGKEVTESILGEIFSKFCIGK

949230 950159 PAUD 949230 950159
psdD-Phosphatidylserine Decarboxylase
FLFIVSRGLVQKPQYIDRITKKKVIEPIFYEKTMLFLYNSKLGKKLSVFLSTHPIFSRIY
GWLQRCSWTRRQIRPFMNRYKISEKELTKPVADFTSFNDFFTRKLKPEARPIVGGKEVFI
FPVDGRYLVYPNVSEFDKFIVKSKAFSLPKLLGDHELTKLYAMGSIVFARLAPFDYHRFI
FPCDCLPQKTRCVNGALFSVHPLAVKDNFILFCENKRTVTVLETEQFGNVLYLEVGAMNV GSIVQTFSPNQTYAKGDEKGFFAFGGSTVILLFLPNAIRFDNDLLKNSRMGFETRCLMGQ

CPn_0840 950141 951544
CT700 hypothetical protein
ISERRNLKTLKTFFGIAKRDKSOKWRIMMLVILWALAASLAIALVAKGYYRFVYFRRYAV
QVIREVRLSMELKEWALAEQQLLPILKKRSYRRQCLFEYMRILKKMORFEESEKLLAEAK
KLGLRGPYFFLEIAYKAYRFGAFKECAQAFASVPQDLFEEEDAAKYASALVRLGDLDAAC
SLIEPWISPLSHQETFVTMGHIYFTSKRYKDAIDFYNRANALGVCPVEVTYNLAQAYRIT
SSYAKAGKLFRKLLSNEVYKEALFNIGLCEQKLGREPGKALLIYQSSDLWSKGDALLMKY
AAMAAMDQRDVVLAEPCWELALRCSTFAKDYKGGLGYGFSLCRLRKYGDAERVYCNLIQN
FPECLTACKALAWLCGVGYATLLGSEEGLMYAKKAVELDHSCETLELLSACEARCGNFDA
AYEIQSFLSSRDTSLQEKQRRSQILRILRKKLPLNDHHIVEVDALLAA

951719 954640 SecA-Translocase SecA
IKRHMLGFLKRFFGSSQERILKKFQKLVDKVNIYDEMLTPLSDDELRNKTAELKQRYQNG
ESLDSMLPEAYGVVAVVCRRLAGTPVEVSGYHORWDMVPYDVQILGAIAMHKGFITEMQT
GEGKTLTAVMPLYLNALTGKPVHLVTVNDYLAQRDCEWYGSVLRWLGLTTGVLVSGTLLE GEGKTLTAVMPLYLNALTGKPVHLVTVNDYLAORDCEWVGSVLRWLGLTTGVLVSGTLLE
KRKKIYQCDVVYGTASEFGFDYLRINSI LATRLEEOVGRGYYFAI IDEVDSIL IDEARTPL
IISGFGEKHNPVYFELKEKVASLVYLQKELCSFIALEARRGLDSFLDVD ILPKDKKVLEG
ISEFCRSLWLVSKGMPLNRVLRRVREHPDLRAMIDKWDVYYHAEQNKEESLERLSELYII
VDEHNNDFELTDKGMQOWVEYAGGSTEEFVMMDMCHEYALIENDETLSPADKINKKIAIS
EEDTLPKARAHGLRQLLRAQLLMERDVDY IVNDDDIVIIDEHTGREPOGRFSEGLHQL
EAKEHVTIRKESGTLATVTLQNFFRLYEKLAGMTGTAITESREFKEIYNLYVLQVPFKP
CLRIDHIDEFYMTEREKYHAIVNEIATITIGKGNPILVTTESVEVSEKLSRILRONRIEHT
VLNAKIHAQEAEIIAGAGKLGAVTVATTMAGRGTDIKLDNEAVIVGGLHVIGTTRHQSRR
IDRQLEGRCARLGDGGAAKFFLSFEDRLMRLFASPKLNTLIRHFRPPEGEAMSDPMFNRL
IETAQKRVGGRNYTIRKHTLEYDDVMNKQRQAIYAFRHDVLHAESVFDLAKEILCHVSLM
VASLVMSDRQFKGMTLPNLEBWITSSFPIALNIEELRQLKDTDSIAEKIAELLQFQK
PDHMVEGLSKAGGEELDASAICRDVVRSVMVMHIDEQWRIHLVDMDLLRSEVGLRTVCQK
PDLLEFFKHESFLLFESLIRDIRITIARHLFRLELTVEPNPRVNNVIPTVATSFHNNVNYG
PLELTT/TDSEDQD PLELT//TDSEDOD

CPn_0842 955015 954710 CT702 hypothetical protein (frame-shift with 0843) KYYTPPTISRSPWSNIALKTISEPEYDCNOLLKTOSLLTTHVDTLLNAPKDFPNSKNOKH LLFCIAINTLSHYAQFLIAGNRRKFWIRYYNDOVWSEWTPFI

CPn_684: **843** 955330 954994 CT702 typothefical protein (frame-shift with 0843) HKRELICOMREKVENTION/INTENTIONOFFEDELENGERASCEDKYOETOVYVEENTDQC DEPTATOESAŁKWI UIOEO

CPO_0844 847 956751 CIGU-0844 - 3-4 1 - 956791 - 957570

TOVARRADGROEED HETYKIS INDIVITIONAL DIPOSITEEDAHLAKLELINAROVPAQV

IDMYSTORIN:EDYFOKHUNGAGGIENHIJ KREIJA IVNGOEYTTROUTAGELIIAFGVPAQV

KING ETTIMLK I'A ILGARAVGROGUENHIJ KREIJA IVNGOEYTTROUTAGELIIAFGVPAQVI I'A ILGARAVGROGUENHIJ KREIJA ILGARAVGROGUENHIJ KREIJA I'A ILGARAVGROGUENHIJ KREIJA I'A ILGARAVGROGUENHIJ KREIJA I'A ILGARAVGROGUENHIJA I'A I'A ILGARAVGROGUENHIJA I CILEBLE/VDEHEEGEAALPONTFPDFDEVFTEGFOPEEF
NVGKOSIINGLLNEERCIIONTFOTTROHIDILYJHKD IDTAGLRKMKSVKNSIE
WIGGGRTEKA 138ADICLLV IDATOKLUGYEKRILGLIG......PHIILINKWDLLEEVRM
EHYCKDLRATDPYL/QAKMLCISATTKRNLKKIFJA IDELHHVVSNKVPTPIVNKTLASA
LHRNHPOVIGGRERIYYA IQKTTTPLOFILFINAKSLLTKHYEYYLKVTLKSSFNLYGI PFDLEFKEKPKRIIN

158151 156850 CPn_0845

CPD_0845

SHI51 35850

CDD_dy A Edymetric Alexandry CFLENW WOAVE MERCYDDMLMN

RPLEDIDIATNASPTIVSTIFPDVISIGNAGII MARCANDRY CFLENW WOAVE MERCYDDMLMN

RPLEDIDIATNASPTIVSTIFPDVISIGNAGII MARCANDRY EVATERSDEYKDGHP

RIIFSSMREDALRROFTVNGMYYDFEDKVFDFVEGTRDIEKKVIRAIGHPRLRFSEDK

LRILRAIRFSSSIGFTLDPTTERAIIKEAPALVNSVSPERIWQELKKWILKROPYGALSLL

LKLKVLIFIPPELRDI PYSLLRTTIEFARKFNPTHF PEILFILPLFQGVSEBAATVAFGR

RISNRELKIESWYSALPHFONGSGNRVFWAHFLASPTAPLFLELFSALGKDPSRQHF

ISRVQELESRLEGFILRIKTSSPVVSAPDLIAKGISPGRLLGDLLREAEILSIENECLDK EKILLLLOEKGFWK

959383 958112 CPn 0846

CPL_0846 959183 958112
ClpX-CLP Protease ATPase
REHMNKKNLTICSFCGRSEKDVEKLIAGPSVYICDYCIKLCSGILDKKPSSTISSAPVSE
TPSOPSOLRVLTPKEIKKHIDEVVIGQERAKKTIAVAVYNHYKRIRALLINKQVSYGKSN
VLLLGPTGSGKTLIAKTLAKILDVPFTIADATTLEAGVVGEDVENIVLRILQAADVTDVA
RAERGIIYIDEIDKIGRTTANVSITRDVSGEGVQOALLKIVEGTTANVPPKGGRKHPNQE
YIRWITENILFIVGGAFVNLDKIIAKRIGKTTIGFSDDQADLSQKTRDHLLAKVETEDLI
AFGMIPEFVGRFNCIVNCEELSLDELVAILTEPTTNAIVKQTMELFAEENVKLVFKKEALY
AIAKKAKQAKTGARALGMILENLLRDLMFEIPSDPTVEAIHIQEDTIAENKAPIIIRRTP

960019 959387

CYM_US47 350017 35356 C1PP-CLP Protease Subunit KLFDEETOMTLVPYVVEDTGRGERAMDIYSRLLKDRIVMIGQEITEPLANTVIAQLLFLM SEDPKKDIQIFINSPGGYITAGLAIYDIIRFLGCDVNTYCIGQAASMGALLLSAGTKGKR HALPHSRMMIHOPSGGIIGTSADIQLQAAEILTLKKHLANILSECTGQPVEKIIEDSERD FFMGAEEAISYGLIDKVVTSAKETNKDTSST

CPn_0848 961556 960177
tig/murl-Trigger Factor-peptidyl-prolyl isomerase
v0ASSPAPPFKSNKKGCLVPRSLSNE0FSVDLEESFGCIVSALVKVSPEVLNKLNKQALK
KIKKEITLPGFRKGKAPDDVIASRYPTTVYRKELGELVTODAYHALSTVGDRRPLSPKAVR
SNSHTOFDLQEGAKVEFSYEAFPAISDLPWENLSLPQEEAASEISDSDIEKGLTNIGMFF
ATKFFVERPSOEGDFISISLHVSKSNDENASSAAIFENKYFKLSEEMTDAFKEKFLGIS
TGHRAVETITSPEIGSFLRGDILTFTVNAVIEVSIPEIDDEKARQLQAESLDDLKAKLRI
QLEMOAKOLQKRFSEAEDALAMLVDFELPTSLLEERISLITREKLLMARLIQYCSDEE
LEKRKSELIKEAEEDATKALKLLFLTHKIFSDEKLTISREELQYMMDVCSRERFGQQPPK
DISNTTLQELVMSARDRLTYSKAIEHVLRKAELLASTPSA 960177 961556

THE THE TOTAL OF THE PROPERTY LPYEDEYAEIANYHLLNSFYTLPYYDEVRAICDMSYLDOELEAKLHFLYGSLRVPAASLA
LOYODVARFISDOEILARNLVEERKMLEEVFSGFIYDERDGAFRVKSEKKIVEFMTETIP
ANGHBITFNCPENLSGOFIYDETIFELSFREGSDINYYEADLKVHGLLKGVPLDLLWDCI
SAKRFLELPKAGQOSKGTRRGKVNSGKLPCILVLDLEKIAPVVOJIFNEIGFKVLDDLVC
KCHMBLTGISLDOFEALPVNFSMSERLIEIOKOIRGEIEFDFDOVPQOJOATLRSYOTE
GVHWLERLRKMHLNGILADDMGLGKTLQAIIAVTQSKLEKGSGCSLIVCPTSLVYNWKEE
FRIFMPEFRTLVIDGVPSQRRKQLTALADRDVAITSYNLLQKUVELYKSFRFDYVVLDEA
HIIMMTRNAKSVKMIGSDHRLITGTPIENSLEELWSLFDFLMFGLLSSYDRFYGKYI
RTGHEMGNKADNMVALKKKVSFFILRRMKEDVLKDLPPVSEILYHCHLTESQKELYGSYA RICHEMONAUMVALKANSFILKRANEUVENLEPVSELLINGHELESVELLISTA ASAMSELSRLVKQEGFERTHINVLATUTELKGICCHPAIFAKDAEPGDSAKYDMLMDLU SSLÜMESCHKTVVFSQYTKMLGIIKKDLESRGIPFVYLDGSTKNRLDLVNQFNEDPSLLVF LISLKAGGTCLNLVGADTVIHYDMWNPAVENQATDRVHRIGQSRSVSSYKLVTLNTIE KILTLQNRKKSLVKKVINSDDEVVSKLTWEEVLELLQI

CPn_0850 965254 966390
mreB-Rod Shape Protein-Sugar Kinase
LGKKYWNCCRYDFMSPHRNLFKLKNFSNRLYNRALGRFDKVFNFFSGNVGIDLGTANTLV
YVRGRGIVLSEPSVAVDAQTHAVLAVGKKAKAMIGKTPRKIMAVRPMKDGVIADFEIAE
GMLKALIKRVTPSRSVFRPRILIAVPSGITGVEKRAVEDSALHAGAQEVILIEFPMAAAI
GVDLPVHEPAASMIIDICGGTTEIAIISLGGIVESRSLRIAGDEFDECIINYMRRTYNLM
IGPRTAEEIKITIGSAYPLJDQELEMEVRGROOVAGLPITKRINSVEIRECLAEPIQQII
ECVRLTLEKCPPELSADLVERGMVLAGGGALIKGLDKALSKNTGLSVITAPMPLLAVCLG
TOKALEHLDGFKERKGNIV TCKALEHLDQFKKRKGNLV

CPn_0851 966378 968195
pckA-Phosphoenolpytuate Carboxykinase
REFGIVMWSTNIKHEGLKSWIDEVAKLTTPKDIRLCDGSDTEYDELGTLMESTGTMIRL
NPEPHPNCFLVRSSADDVARVEOFTFICTSTEAEAGPTNNMRDPGEMRELHQLFRGCMO
GRTLXILVPEPCMSPLDSPFSIVGWELTDSPYVVCSMKIMTRMRGDDVLKSLGTSGKFLKCHH
SVGKPLSPGEADVSWPCNPKSMRIVHFQDDSSVMSFGSGYGGNALLGKKCVALRLASYMA
KSOGWLAEHML I IGITNPEGKKKYFSASF PSACGKTNLAMLMPKLFGGKIEC IGDDIAWI
REGROGRIZVANNEYGFFGVAPGTGSRTNPNALATCRSNSIETHYMLTAADDVWMEGLTE
QPEPELTDWLSKFWKPGGSPAAHPNSRFTAPLRQCPSLDPEMNFOGVPLDAIIFGGRS
ETIFLLVFALDWERGVTJGAGMSSTTTTAAIVGGLKLEHDEPFAMLPFCGYMMAYYFGHWL
SPAENBCLKLEK FGVNWFRKNNGGSFLWPGFSENLPYLEWIFGTDGLEDIAETTPIGY
LIPHIQKFNLWGLNLDLQTVQELFSVDAEGWLAEVENTGEYLLIFGSDCPQXITDELLRIK
SEELKEK CPn 0851 966378 968195

968274 970513

CTT_UNS2 968274 970513
CTT/TT hypothetical protein
TKELTDYYLINTVTLOPSYINFTPNVTTALSOCKILISATELSCSALFFOELODKAOG
TKHALTDVOELSAKALRYAQVOTSISYLPTEESSRYSJOAGIIDRIMPTPTDDEVKAILQ
NEMPETSKIPVEGLDKVFKSYLDSVTPPEGIDPSNPYSATILINYTTLLNNLKPKFAASST
PTHADVHALYALRIDPVKETEALKAADAPPSKVHAFVSCHIMTYNNMOVLSYPVTDYLN
VQTABLSSATITAAGEVOGYLKNEYSILKDILHRSV DEGATHYPADAEYNARDAGVIQSL

LNLSGNYROLTENMLPHTE I LAQUES FOR VIVAGTI LASHTUL PTTMEDTLIGV
LYTYQCCATIFOMSYGTSTP IDALNQEKSYMOARANG POVTODOVPODFATNIQS
GTSYRGIOLFKINKVNELIP LFLIQAASFLEK PYNLMSRSMYQT LEDAANRS ITALDGLI
SGWSTO LATFOTOKNSLOPSLLKYFOTMKANKESFVTTAPLOMYYSSLMLDKYLPTQNIV
LASLGIOMTYSNKAAKYLIELI KEITTFOSAD LYYSLSTLKKOMLOAVADP IGKAVGVU
INDEKTRAMADITRCNKIKAAI DIMLVEIKADAELSKSQIRELVOTLTNFKSQSDDLIRNL
SCLLGFLSGLTLKAVNDPHATYEAFTAE FTEPFNNWKROLATFESFVIQGONGITPGG
OOQLLOAMESSQODFSTPICNOQUALCYESSAMCOEWTLVSAALALLINGMYSKIARRIKS

470637 971803

CTD_USC 1
CTT12 hypothetica: protein
NIMPKIEKRNSLPLTAVAPVFEES/HPSVATTVDYVDATTLSRHLTVLKDVIKEARNLD
LGKAFLTSMKOGF INTOTELAI IQASLADQSSRESRKEEKI FHQHLGKAAPQAATATSG
VOPTADPVADKMPLOSAFAYVLLUKY IPAQEEALYALCREINLSGYAQNLFSPLLDMIKS
FNSAPINYNLGSYISQTSGTANFAYGYEMILSRYNNEVSQCRLDIASTVKAKAALANNSA
SVKANVSLTDACKKQIEDI IAST KKSLDVIHTQLTDWINLASITFVPGLNKYDPSYRIV
GGDLSI IALQNDEKVLVDGKVØITTAVNEGGLLNFFTTVLTDVQNYGDLAQTQQLMLDLE
LKAMQQQWSLVSASLKLLNGWTTVISGFKN

9/2849 971806 CPH_0854 972849 971806

ompB-Outer Membrane Protein B

GPFDMNSKMLKHLRLATLSFSMFFGIVSSPAVYALGAGNPAAPVLPGVNPEQTGWCAFQL

CNSYDLFAALAGSLKFGF/GDYVFSESAHITNVPVITSVTTSGTGTTPTITSTTKNVDFD

LNNSSISSSCVFATIALQETSPAATPLLDIAFTARVGGLKGYYRLPLANAYRDFTSNPLNA

ESEVTDGLIEVQSDYGLWGLSLQKVLWKDGYSFGVSADVRHGSSPINYIIVYNKANPE

IYPDATDGNLSYKEW\$ASIGISTYLNDYVLPYASVSIGNTSRKAPSDSFTELEKQFTNFK

FKIRKITNFDRVNFGFGTTCCISNNFYYSVEGRWGYQRAINITSGLQF

974001 CPn 0855 CPT_UB55 974U1 97294
gpdA-Glycerof-3-P Dehydrogenase
GLMKQHIGYLGHGIMGFCLASLLANKGYPVVAWSRNPDLIKQLQEERRHPLAPNVVISPN
LSFTTDMKEAINARMIVEGVISAGIRPVAEQLKQITDLSVPFVITSKGIEQNTGLLLSE
IMLEVLGDSVTPYLGYLSGPSIAKEVINGSPCSVVVSAYDSQTLKQIHEAFSLPFFRVYP
NTDIKGAALGEALKNVIATACCIAEGLSFGNNAKAGLVTRGLHEARKLAAIMDCKPETLN GLAGIGDLC/TCFSESSRNLRFGHLLAQGLTFEQAKAKIGMVVEGAYTALSAYQVAKHHK IDMPITTG/YRVLYENLDLKEGIALLLQRNTKEEFL

CPn_0856 975410 973995

AgX-1 Acmolog-UDP-Glucose Pyrophosphorylase
GSRDRVVRLTVMTESVYSPSAMHVNSLADKLKAINQEHILDIWPSLSPKQQQRLFQQLTS
VDIDFFRKQQQLLSSPTAILKDFHPITSFASSGEDPERAHAGTTLLKEKKVACVVLAGGQ
GSRLKCDGPKGLFPVSPIKKKPLFQLVAEKVRAASKLAGQPLPLAFMTSPLNTRQTRSFF
ESNJYFHLDPNQVDFFCQPLWPLLTLSGDLFLEDMDTLALGPNGNGCIATLLYTSGYWEK
WKNAGIENVSVIPINDPLALPFDVELCGFHAMSNNEVTIKAALRQTAIEDVGILVKSHDS
GKTSVIEVSEIPQNERFALNEDGKLKYCLANIGLYCLSMDFIRHAAYQQLPLYKVHKHAK
QVGHTSLNEKNAWKFEEFIFDLFCYSDHCQTLVYPRQECFAPLKNLEGNNSPDTVRQALS
DRERQLFHKVTGKKLSPNTTFELEADFYYPSTSTSLHWENKAFFEEPFFEAS

975392 975808

CT7]6 hypothetical protein
LLIARQYIKTARGISRLMRDRLGSLSLILKVKIHKYLDTLHNQKRLALTVSRNIQATNKR iaólhleryehfisronikhydilleylktlosslykooseslrfleihhooloelinrr Kiiekiknnkyskooeigt

CPn_0858 977115 975757

flii-Flagellum-specific ATP Synthase
RNSETRNQRRTRPSTFCFDSMNHLMKEKLHIHNWQPYRACGLLSKVSGNLIEVDGLSACL
GELCKISSTKDPNLLAEVIGFHNHTTLLMSLSPLHSVALGTEVLPLRPPPSLHLSDHLLG
RVLDAFGNPIDKKEDLPKTHRKPLLSLPPSPMMROPIDQIFPTGIKAIDAFLTLGKGQRI
GVFSEPGSGKSSLLSAIALGSKSTINVIALIGERGREVREYIEKHSNALKQORTIIIAAP
AHETAPPKVIAGRAAMTIAEYFREGGHEVLFIMDSLSRWIAALDEVALARGETLSAHQYA ASVFHRYSETTERAGNIDKGSITALYATLYYATKHPDIFTDYLKSLLDGHFFLTSGGKALA SPPIDILSSLSRSAQALALPHHYAAAERLRSLLKVYNEALDIIHLGAYTPGQDEELDKAV KLLPSIKAFLAQPLSSYCYLDNTLKQLEALADS

977597 977055 CPn 0859

CPT_UB39 97/397
CT718 hypothetical protein
VFLVTTPQSFGSLSQSHLPHPHDPWDTEPTSLPEDPNDKASQELHSLVHLFRKLSIHLLS
EVEKTVQQLKPDLLELALLICEKFLYKKLENPQELALLLSTALQRHTTLRSLTPIKVFLH
PEDLKTLTDWISTHELPMIKHAEFFPDTSCRRSGFKIETPNGILRQEISEELDHLLSVLT

CPn_0860 978639 977608
flif-flagellar M-Ring Protein
RTLVFFONLAKKLTALGISFLGCLLIGGVVSCAILFGRSSNPSLAPTQVKTEKTSGNWLK
LTOMGNPKLIESLTKKEÇLEKDLTSFHPIASAKVAIALSTEDDVMSPLHLSVILTLRKEE 978639 977608 SLTPSLIFSITDYLCSSLFGLKREHISLSDNLGNLYIPESITVNSLFIHTLENYLGKIFP KEHFALAYHAKAEKPTLQLTLNENYIAHLTKEESEKIVAHTKHYLYQNYDDSYDIVIETL PFARLQNKKSFPAKVLIGSMILVISLMIVALASFYLARHAYERVSPEPRKIKRGINISKL LEIIQKESPEKIALILSYLDPKKAEALLNRLPEDLKHQVLKYKL

CPn_0861 979752 978925 nifU-NifU-related protein ASYPTWARFLMTLPHEPMIFWSSLSAK/MKKFLTPHCAGTFSEEDARAKEAHLVTGKOGH ASTPT WAT BUT TELEPT TO SUSSAIN WARE LIFTLAND TO SEED ARRANGED VINOUR REMONCYTEYMUZDKKNOY ILDAKEQYECHEYLIFLAEAVCHLVCGKSYSSAYKHTLDDI DKSLRYHAHQPALPEDSITSLYHFVIDALDTAVEQCLETPLEEGSLPLUNSPMNLDFEDAN PYSOSUMEALTHEOKUZALRATTAKKIOPYIAMDGEVTYESLENFIVTIAYSGNCSGCP SSLGSTLNSIGQLLRAYINFELQVKVDESSLNLCHP

080924 ytho-mits-related protein GROTIFRITDOKTROIRMENTONRKAPPIPWLNNOVAIPPSERVKERYALHRDIFSLPPG GROTTE PETELOTAL CHREN GYIRKARP LEMONIOVA I PPOERWEST ALIEUTE DEPO SALKLAEKTEES TROLUCERCHI EREF ZHEFHEIVANI I VI.AALVEELSHEG IRRIH I LEPH DOOLL INSLORHOOL TTY CWYTVNHEYD LVEEQL I ETLEPESILLESII SAARGETOV TOP LOPELD LOKORRE I LIHILDESIO DERRAFLETEEL LAID LETTETISTINE KELL GOEL STUDIES GLAESEVONREPHIVAAL PED PAGEILAEFILHIGAS I VICILIY EREPOLAOVLON GEOFF GLAESEVONREPHIVAAL POLIFAGEILAEFILHIGAS I VICILIY EREPOLAOVLON GEOFF LCHOALHEST TEROKOLEFOKLARAMIDA EKHLT PLLCCOS

CPn_0863 Q8/\$20981520 980831 pgmA Phosphoglycetate Mutace 980831 EHMALLILLEHGOSVWNEKNLFSGWVDIPLSQQCIEE AIQNLPIDCIFTSTLVR SLMTALLAMTNHHSKKIPYIVHEDPKAKEMSRIYSAE PLYQSSALMERNYGELQ GKNKKQTAEOFGEERVKLWRRSYKTAPPQGESLYDTKQN—EYYFEKNILPQLQNGKNVFV SAHGNSLRSLIMDLEKLSEEEVLSLELPTGKPVVYQWKNHKIEKHPEFFG

CPn_0864 981658 982374
yjbC-predicted pseudouridine synthase
yGVNYTKVRLNKELASATVASRRKCDEIIFSGSVTVNGRVAEGPFVLVDPEDKVOVGGTS
YHDERGYTHMAIRAUTH. LEHEFFFFF THELYTDLFAHLEYRVETVGFLDKETDGLLLVTH
DEFEARKITHESSUTTESSLENZSPPSSAKDURTMESTEIDSEINFFZDYTTEPPTTVK
IVVSEGKKHEIRLFADAAGFPILELKRIRIGSLVLGGLRYGEYRELTDAELGTYMKLSD

CPn_0865 982412 982942
CT865 hypothetical protein
SPMGYVFYVIAGSIFLGISLGAYCQLYYSVKSVLFSWYLLTVYALEKRHALLALSQLVGE
EDAQSQKEIDFLSQCDKLSWRAFLKNSYEIIPTFKEMEDLLSERVQGFLESIETIAEHDR
AILCIENFWASKNLFDFEIAAYEEAVEKYLKLRQRAPLRLASKLFRFLDVPSIRFSS

CPn_0866 983494 982916
bira-Biotin Synthetase
mkWIYYEIEEIPSTNTMAKSYMHLWDPYALTVISTKCQTAGTGKFGKSWKSSKGDLLNT
FCFFITDLHIDVSRLFRLGTEAVVALCKDLGITEAKIKWPNDVLVHGEKLCGVLPETLFV
EGLLGVVLGIGLKGNTTKQALKDVGQPATSLQEILGHPIDLETTRELLIHHLLGVLQENL
PDSLATKSNRGNI

CPn_0867 983405 984667
rodA-Rod Shape Protein
CIRIPOMHIGFCHCVRGGNFFYFVINNFHILEIYSLLNSNTIMRYHKYFRYVNSWVFLVV
LITLMLLSVVVISSMDPTAMLVTSSKGLLTNKSIMQLRHFALGWVVFFICAYFDYHLFKRW
AWVLYFFMICALVGLFFVBSVQNVHRWYRIPFIHMSVQPSEYGKLVIVIMLSYILESRKA
DITSKTTAFLACLVVALPFFLILKEPDLGTALVLCPVTLTIFYLSNVHSLLVKFCTVVAT
GIIGSLLIFSGIVSRQKVKPYALKVIKEYQYERLSPSNHHQRASLISIGLGGIRGRGWK
TGEFAGRGWLPYGYTDSVFSALGEEFGLLGLLFTLGLFYCLICFGCRTVAVATDDFGKLL
AAGITVYLAMHVLINISMMCGLLPITGVPLILISYGGSSVISTMASLGVLQSIYSHRFAK

CPn_0868 986733 984670

ZDLA/CAGA-METAI Transport P-type ATPase
NFRNGLGVRDLHHFREYYLIINEIIITGRYVFSRLFFTSFSAEVUNTFFESGMSEDTSPL
LSKORKKLSHNLPLKSAYLSLGTYLIALLSFWLHAKNLSNLFVVPTFFLAGTPALIKSLD
NICOKVVNIDIILMTSAAFGSIFIGGALEGALLUVLFAISEALGOWYSKAKSTLVSLKQL
APTTGALVLEDNILOKVAINKIEVGWILRIKSGEVVPLDGETLHGSSSINLHHLTGEKVP
KSEHPOSIVPAGAHMMEGSFDLRVLRTGSDSTIAHIINLVIQAQNSKPRLQORLDKYSSV
YALSIFAIACGIALLVPLFTSIPLLGPQSAFYRALAFILIASPCALIIAIPITAYLSAINA
CANHGYLLKGGVILDRLVSCNSVVMDKTGTLTTGELTCIGCDYFGSKNETFFFSVLALEQ
SSSHPIAEAIVSYLMEQKVSSLPADRYLTVPGEGVRGYPNEQEAFVGRVETGLGKVPSEY
LEDTEQKIVQAKOHGEICSLAYVGNSFALFYFRDIPROPAKRIIQDLKDUGYVSMLTDD
HKVSAENTAEILGISEVFFDLTPEDKLAKIRELATORQIMMVGDGINDAPALAQATVGTA
KGEGSATAIEADIVLLHDSLSSLPWIIQKAKQTKKVVSQNLALALAIILLVSWPASLG

CPP=0869 987479 986658
CT728 hypothetical protein
EGM#FFFPKTSENTSDCROHQILRKIMTQDPHDHFKSRTPEDHIKHVRDKHRVCKGEPHT
FFKFFFPKHANNALSTOV IFFIRTHFFLIPTNRALQVKSLISLGVGWTFYHGCLKARKA
WAYMELSHRSMLEEKWEIEENFEQEKIELRILFENGGFKDPLLQEWEFVCSDSTLLLDT
MIREELYIRKEDLPPHPLIQGGSRILGGLCGLAIFLPLVLCISYTLAGVFSALMVLVLSFL
KAKILKNDKISEMVWVLGIFITSASIISSLMKLL

CPH-0870 988881 987448
set&F.Seryl trna Synthetase-2
TTHPTOGFGGAVILPFSPISIARRIKKSCCSEKSSIYSHFCTLLLNNETSMLDIKIIRK
TP\$ECETRLRKKDPKISLEFVLSLDKEVRQLKTDSETLQAQRRLLSQDIHKAKTQGVDAY
NL 10EVETLAADLEKIEQHLDOKNAQLHELSHLENYPADDIFVSEDDAGNOVIKSVGFL
PIE\$FPFRKHLELNQELDILDFQAAAKTTGSCWPAYKNRGVLLEWALLTYMLQKQAAJGF
QLWLPPLLVKKEILFGSQIPKFDGQYYRVEDGEQYLVLIPTAEVVLAGFRSQDILDEKE
LPBEYAACTFCFREAGAAGAGERGLVRVHQFHKVEMFAFTTPNQDDIAYEKMLSIVEEM
LTELKLPFKISLLSTGDMSFTASKTIDAEVWLPGQKAFYEVSSISQCTDFGSRRSYTRYK
DSGGRLQFVHTLNGSGLATPRLLVAILENNQQADGSVVIPEVLRPYLGGLEILLJKDQ

CPn_0871 988766 989899

ribD-Riboflavin Deaminase
EYMEDFSEQQLFFMRRAIEIGEKGRITAPPNPWVGCVVVQENRIIGEGFHAYAGGPHAEE
LAIQNASMPISGSDVYVSLEPCSHFGSCPPCANLLIKHKVSRVFVALVDPVPKVAGQGIA
MLROAGIQVYVGIGESEAQASLQPYLYQRTHNFPWTILKSAASVVGGVAJSQGKSQUHTC
PEAPHDVGKLRAESQAILVGSSRTVLSDDPWLTARGOPCOMLYPKQPLRVVJDSRGSVPPTS
KVFDKTSPTLYVTTERCPENYIKVLDSLDVPVLLTESTPSGVDLHKVVFYLAQKKILQVL
VEGGTTLHTSLLKERFVNSLVLYSGPMILGDQKRPLVGVLGNLLESAPPLTLKSSQILGN
SLKVVWEISPQVFEPIRN

CPn_0872 989903 991216
ribaribb-GTP Cyclohydratase & DHBP Synthase
kEPIFRVACLASESVNARESMIETREEVGSANFVSLEPAIEDLEAGKFVIVVDEASREDE
GDLIIAGEKTVEKMTFLLQHTTGVVCAALSQERLLSLDLPPMYKDNRCRFKTPFTVSVD
AAHGVTTGVSAADRTKVVQLLADPKSKPEDFISPGHFFPLASEPGGVLKRAGHTESTVDL
MELAGLQPCGVLAELVNEDYSMMRLPQILEFARKHNIAVIPYTSIIAHRMLSDRLVSKIS
SAPLPTIYGDFTHVVESLLEGMQHLALVKGNVAGKSNVLYSVHSECVTGDILGSKRCDC
GEQLSSAMSYIAEKGTGVLVYLRGQEGRGIGLGHKVFAYALQDNGYDTVDANLAMGFFVD
SREYGIGAQILVDLKLTTIKLITHNPQKYFGLQGFGLSIFERVPLPVRISEDNEQYLRTK
QEFMGHHUDLFCCNNRVQ

CPG_0873 991188 991694
cibE-Ribityllumazine Synthase
LSHAYTIGYMNEESYMTIKGHIBAKNIRIA IVOSEFIJAMADALVAATQETELKEGGSE
ISHAYTIGYMNEESYMTIKKHISHEKEDA IVACQVLIGGETHIMNQIVNQVAAGIGALS
LEECLPITTI IVAADSAE IAWQRSGIKGRHIJVSCMTTA IEMATLETQI

CPT_0874 993164 991749 CTT733 hypothetical protein LICOLINK LURKONDRERASMEK LEKT KYLVYPLALLMOCHO (GYAGINNOLOTKVK TOSETWI POKLROYPELLMITESOGAPULOTTP IDMAYSEKLFNKKYPALDIATRSMIHL HLLIQUISROSYMOLIQUILE THROEOTAHKOLLEFUNDEKSEENTUR ILLETAL HEROEOTAHKOLLEFUNDEKSEENTUR ILLETAL ILLETAL BETYAKALHVURTETELLOTUVARLIGEQQUEVLLSLERL GNYDSLINLTEYPSAQLLIGAMRTRISLAILOLYYVOLDTCUDKNCGGEFYINFAPLLISMI GNYDSLINLTEYPSAQLLIGAMRTRISLAILOLYYVOLDTCUDKNCGGEFYINFAPLLISMI GNYBALLIGARSYTTARARLIGEGTSRTDMTLVRIATLMGLIGGEASTLAWSFKNLPSD GAENLVNSFYTVQGEHIPLTFRGLPSLVAGLSVATHGSTVSPENRLRQLYSTMLSLLVKSLRSHREMLIKQLLPQCTVLDFSETTLSSGGLVVFAESIAVRIHLNGAVSINL

CPn_0875 993363 994022
LET MERVITY FOR THE STATE OF THE STATE OF FOR THE STATE OF T

CPn_0876 994123 995517
daga-D-Alanine/Glycine Permease
SIATOETMILYFIEQLINKLSTSFOVFPHILLIGGFLTWKLRGLQFHGLKLGFNLMLQNKLD
DSSSKANEVSSYEAVAGILAGUFGTGNIAGMAVALACGSGALVWWLAALLGAIVQYAG
SYLGSKYRKPEGNTGEFIGGP/ACLAFGMRKKILAGFFALFTIMTAFCAGNCVQVSCIVL
CAGETFEKLLVGFLLALUV/PVLAGGNNRILRFSARVIPFIAGFYCISCGIILFQHASA
ILPAIKLICSSAFGIKAGLAGIGGYTLSQVISTGINRAVMATDCGSGMYSILQANTKSKN
PVVDGLVTLVPPVIVMVVGITMLVLIVSGAYSGAQGTLMWASAFKNSLGSLGSVIVIL
MAALFGYTTILTWFACAEKSLQYMIPGRRANLWLKAIYVLIIPLGQVIDMRMIWALSDTG
FSGMVILNCIALIALLKDVLSTNRDVALLKERECSVADPVRNLDA

CPn_0878 996660 995992
SET Domain/protein
GCMSTVTTE/CSSIHISIANDWRDSQPYSLDRASELLHFRFLPSLVFSNWKVEQQIETLC
HKSEKRRLISPLAKWLGKLHKQDLLCPPAPPVSVCWINAHVGYGYFARDEIAPMTYIGEY
TGILRHRGATWMDENDYCFRYPMDLFTLRYFTIDSGKGGNVTRFINHSEQPNAEAIGVFS
EGLFHVIARTVAPIYAGQEICYHYGPLYWKHRKKREEFIPEEE

CPn_0979 997463 996645
yycJ_metal dependent hydrolase
yrilwxysmogffplasgskonsaylgtdsckilidlgvskqvvtrellsmidpediqa
ifv/hehsdhisgiksfykayntpivcnletaralchlldshpefkifstossfcfqdle
voyfnvphdavdpvafiffyreekigfctdlgwytswithelydcdylliesnhspelvr
gsdrpdvykkrvlskighisndecogligkiitpklkklylahlstechtaelalstvse
såasitsiapeialaggitspiyfsrlevacpr

CPn_0880 999864 997444

ftsk-cell Division Protein Ftsk

PHIRERKESRHPRLPTLPLAAKASLYLFFACFSGLSLWSFHRDOPCTONWIGLLGWSFSS

ELLYFFGAAAFFIPLYFLWLASLYYFRATPRPLFFYKAAAFLSLPFCSAILLSMLSPVGTL

PALLDTRLPKFILGNNPPVSYVGGIPFYLFYEOGSFCLKHLIGSVGTALITGFVMLFSVT

LYCGGIALLKKKTFODGVKKAFCSFFOTCFKNLKKLINRRNYLPKPSVPFVSKNPFSCTK

SØPSPRRVSETIILDGSISPLPQEEIPGSKKESFFLTPHPCKRFLTKFVEPOENKAKEGK

TIALSSTPTVVRESKGERAALPKLKSLAVPENDLPOYHLLSKNREARPESLOAELERKA

LILKOTLTSFGIDADLGNICSOPTLAAFEVLPHSGVKVQKIKSLENDIALKLQASSIRII
APIPCKAAVGIEIPTPFPQAVNFRDLLEDYQKTNRKLQIPLLGKKANGDIMADLATHI
APIPCKAAVGIEIPTPFPQAVNFRDLLEDYQKTNRKLQIPLLGKKANGDIMADLATHI
SREVYNALJWLVKEMESRYEILRYLGLRNIQAFNSRTRKKTIEASVDREIRETMPFWGI
IDELSDLLSSSQDIETPIIRLAGMARAVGIHLILATORPSREVITGLIKANFPSNISFK
VSNKVNSQIIIDEPGAENLMGGDMLVLLPSVFGTIRAGGAYICDEDINKVIQDLCSRFP
TOYVIPSFHAFDDSDSDNSGEKDPLFAQAKTLILQTGNASTTFLQRKKKIGYARAASLID
QLEERRIIGPSEGARPROILLIONPLEG

CPn_0881 1005646 1006209
No robust homolog present in Genebank/EMBL as of 11/7/98
NKKRAVHMPVPIONSSRNLQEVPESLEDLEOHAEESPTHQSAESSSLQLSLASSAISSRV
EQLSSLVLGMENSDFSSLRDVPIFSAIYESSTHTPVPTPLVGVGYINGSQSGYYDTQRES
LHLSQLLGSRRVEVVYNQGNFMEASLLNLCPRRPRRDSPISLALLELWEAFFLEHPPGS
TFNPIFFW

CPn_0882 1006169 1007404
No robust homolog present in Genebank/EMBL as of 11/7/98
NTPOVALLIQYFFCNGAFYVREALRLTPHAGNIVLVGICPSLYPEHPRSFYYRVSGDIGS
RFDDRGFVNSCVETLPYSSGSFGIFWISFTDPTFNFAIWNTFMRTAGINEVSREMTODTE
TSLIEMROLSEQGEANNTDSLEQEESLMGIVGHTVGGVSMTVTSSPNIFYRIQTLLGLPE
TLAEAEENPTFPNSTIDSLAEIMMNLVRISDAVSIFWIFPIVDTTYNGVLLAVCIGFFGI
NGICSTFLMLTNPRSRRDRWRNLRIMVLCYRSLGSGMULFDLSNNVRMARRHYTSCTVL
LYAMVTLFGWTVAIODALQYGFPSVRDAFYRYCLRHRYCLTQRNEDSLQTTGTRFQVTRT
HLEDQQMVASILNLSVFGLFFGFVGLMTTFGGLEISPSCRWDAANNRTVGIF

CPn_0883 . 1008904 1007573

dmpP/nqr6-Phenolhydrolase/NADH ubiquinone oxidoreductase
LYELFIKSGIFIVMTWLSGLYFICIASLIFCAIGVILLGRKLFIKVHECKLKIND
NEELTKYUVESGOTLLUVELLSSGIFIPSPCGCKATCKÇCKVRYVKNADEPLETDRSTFSKR
QLEEGWRLSCQCKVQHDMSLEIEERYLNASSWEGTVISNDNVATFIKELUVAVDPNKPIP
FKPGGYLQITVPSYKTNSSDWKQTMAPEYYSDWEHFHLEDQVIDNSQLPADSNAKYSLA
SYPAELPTIKRNIRIATPPFINGRNOEIPWCVCSSYVFSLKRGDKITVSGPYGESFMKD
DDRPLIFLICGAGSSFGRSHILDLLLNKHSKREIDLWYGARSLKENIYQEEYENLEROFP
NFIYHLVLSEFLFEDIAAGWKDDPTKTNFLFRAFNLCQLGRLDNPEDYLYYVCGPPLHN
SSILKLEDDFOSSIILDDFGS

CFn_0884 T009368 1009099
CFn_0884 T009368 1009099

CPn_0885 T0 t0650 1009433

yggA cRNA Methyliconeterne
AD LITMOTMODEPHROVERALELEGETZODELKEKEELLIQLEAPLVEOMTAP LITCOP
DEGRIKMEESFECTYFJEKSLAFISTEKKELFUTTERKKGLEVTTSLLITEROTHOLIKETKEMMIKH
PELMAYETEKKOLATITUVETSLEJAJEMOTTSTEFYRNIFACIDEMKETILGOSSI.

NIA:: IYWEEKVAARGISTYYETKILYGADJOOKLULY ASFSLRPRSFFOPQITO AAKIIETAYEF INPEGSETLIDLYGGAGTIGIMLSPY VEI IPDAVASAQENIKA NNKEDLVEVYLEDAKAFCKRNENKAAPDVIII LPPREGR. YLKYILRIGSPKIVYISC NPKTOFOECADLISGGYRIKKMOPIDQFPYSTHLENIILLEREIDP

CPn_0886 1011288 1010908 hcta-Histone-Like Developmental Protein RTLFMALKDTAKKMKDLLDSTOHDLAKAEKCHKAAAORVRTDSIKLEKVAKLYRKESIKA EKIRLBERFI TRALAFYEKTAEFFARFILLAAAARTLEAYEADPPALEFTAAKKVEEGG

CPT_0887 1011692 1014157
CHITR possible phosphoptotein
MKKLYHPTLFLRELIRLSLIFALSUTLISGNFPQQKSFGHCCADMHSALISGKNCEELFA
MKKLYHPTLFLRELIRLSLIFALSUTLISGNFPQQKSFGHCCADMHSALISGKNCEELFA
DFIERVLADRETLTARDWGTVVVLVREYLLKCIRKGDCDYGVKILQKLLALRLPKDARKD
LQILWHRLNPEQAPLRBOVYDOLFTIGCHESLQDHLLFELYTUTHSGYERKKODMLLAKE
GDYKKAIELAKELVAALEKGSCSPHPEIVOIEKTFLOKTLLALQIKVAQEAQESCDALL
TPYCLSEIAYTEAMDALVLRIARGEVSRTNEVDSVLLSHALQHLPFAREKAIPELEVLID
HGAYLESTLLYYAYFSLLELYHONNDFASLERLLEKGDAVFVPEPHYFPEYGFFLGAYFY
KAGKYESSAENVFLQIIDPAVKLGATFARAYEYLCCIAYVQNHYEKAEEYFLRAYKSVGRE
ESGIGLFLAYAVQKKKTACEDHLYHPKFSFTYRHLLDSLCSLSYPHGENKGSSAIQRVHR
AVPELSEIYSRCIYDMIKYRNTYTHPIIELAYNQVRNLEKRNLEEICRDAQDPEYDKAL
AFWGALQSGASVRSLIESSDVDEARITIRCYEALYFHNDDAIAMLPQAFSEECNSWQTA
LRILWTLVRPKGAPNHAKYWDHLVLRPHGDSLYFFGYDLQEYLIGKEDALKHLSVFAELF
PKSSLLSLVYLQGYSESSALRRVGWFVKALEEFTEISWSGEHMKTWAY IYYMVKLDLAD
TYISLGNFSQAVHILEEVKEDMQVASHPKLHFLKGEDCYLAMELRWVEGLAYAYFQLHET
AHLSNHLLEHVEKNLISPRSYRDYYGESDAARRVGRFLGCOPFLGV

CPn_0888 1015441 1014119
hemG-protoporphyrinogen Oxidase
AERRFCVKRAI IIGAGISGLAAGMUHKKFPQAEILVLDKEAYAGGFVRTESPQGFSFDL
GPKGFLTRGDGEYTLKLIHELGLQNSLIFSDRAAKNRFVYYRGKAHKISTWTLLRKGLLP
SLIKDFRAPCYTQDSSVQDFLKRHSSQNFTSYILDPLITAIRAGHSSILSTHMAFPELAK
REASSGSLLRSYLKNRSPKKSKTDRYLASLSPSMGTLITTIQEKLPATWKFSTSVTHIDC
SPKEACVTTPSETFFADWIYYGPLQQLPVLLPNYGIENLSKRVLPMILSSISLGMHAN
FSLPKGYGMLFADELPLLGIVNNSQIFPQATFGKTVLSLLIEGKWRESEAHAFAIAALSE
YLNINOKPDAFALFSSQDGMPQHAVGFLERKERILPHLPGNLKIVGQNIAGPGLNRCIAS
AYHAICDLHTEETLAQPQSSL

CPn_0889 1016841 1015462
hemN-Coproporphyrinogen III Oxidase
FEMENUNFKFLEGLHOPAPRYTSYPTALEWEPSDAAPALLAFORIRENPOPLSLYFHIPF
COSMCLYCGCSVVLNRREDIVEAYINTLIQEHKLVVETIGFRPQVSRIHFGGGTPSRLSR
ELFFLLFOHHKLFDLSHAEEIAIEVDRSLENDMEKADFFONVGFNRVSLGVQDTQADV
OEAVRROSHEESLKAYEKFKELAFQSINIDLIYGLPKQTKESFSKTIQDILAMYPDRLA
LFSFASVPMIKPHQKAMKASDMPSMEEKFAIYSQSRHLLTKAGYQAIGMDHFSLPHDPLI
LÄFBINKTLIRNFGQYSLPPEEDLLGIGMTSTSFIRGIYLQNAKTLEEYHNVLRGTFATV
KSKILTEDDRIRKWAIHKLMCTFTINKEEFFNLFGYEFDTYFIESRDRLISMETTGLIHN
SEGSEKVTPLGELFVRVIATAFDHYFLNKVSKKECFSASI

CHP-0890 1017829 1016819
heme-Uroporphyrinogen Decarboxylase
STHANWDSMSAFFDLLKSQTASHPPIWLLRQUGRYMPPYQELKGSQSLKTFFHNTEAIVE
STHANWDSMSAFFDLLKSQTASHPPIWLLRQUGRYMPPYQELKGSQSLKTFFHNTEAIVE
ATHLEGPSLLHVDAAILFADILSILDGFAVTYDFAPGPRIQFSPEQPFTFTSDPQTIFSYL
LDARTLKQKLPVPLIVFAASPFTLACYLIDGASKDFSKTMSFLYVYPEKFDQLIFSYL
EGTAIYLKTQMDAGAAAVQLFESSSLRPSALFTRYVTEPNRELIXLKEGALPVSLFGR
CFEENFYTLQATQADTLHPDYHVDLHRIQKNLMLSLQGNLDPAIFLLPQEKLLHYVEAFL
VPLRTYPNFIFNSGHGILPETPLENVQLVVSYVQRQL

CFE-0891 1021079 1017819

mfd-Transcription-Repair Coupling

mfd-Transcription-Repair Coupling

mfd-Transcription-Repair Coupling

mfd-Transcription-Repair Coupling

mfd-Transcription-Repair Coupling

mfd-Transcription-Repair Coupling

LDDL/ENLRTFLDOAPVEFPSSEIDLSPKLVNIDAVGKRDHLLYSLNOHRAPIFCVTTLK

ALJESTRSPOATSQOHLDLAVGDVLDPEATTELKSLGYSQOMLTSEKGEFSFRGGIVDI

FSTEPLYLFDNLELLEDDFADISGTLSSLPDRFFSIGTLYDRISTSNQVYFSETPFPNKL

MLKERRVILEAFHRMEASRQAIPILYPEQIIONDENPLLAFLOHLQETMPHGKPLKLA

TYSTKTKSLKEARALAETVARGDVEIYERTGNLTSSFALVNEAFAAISLSEFASTKVLRR

OKORTHFSVTTEEVFVPI FGETVVHINNGIGKFIGIEKKPNHLNIETDJVLDEYADKARL

YVPSNQAYLISRYVGTSDKAADLHHLINSSKWRSRDLTEKSLIVYAEKLLOLEAQRSTTP

AFVYPPHGESVIKFAETFPYEETPDQLKTIDQIYNDMMSPKLMDRLYGDAGFGKTEVLVLEYADKARL

TYSTCHAMBUSKFARDLSTEFNPOLLTIDDIYNDMMSPKLMDRLYGDAGFGKTEVLY

AFVYPPHGESVITEVFVETPPOLKTIDQIYNDMMSPKLMDRLYGDAGFGKTEVLY

TERTLHMSLSGARDLSVIAMPPLDRLPVSTFWHENHTELTLAALHHELLRGGQAYVHNR

IESIYTLAETIRNLIPEARIGVAHOOMGAEDLSNIFTKFKNOK/DILVATALIENGIDIP

NANTILIDHADMFGMADLYQMKGRVGRWKKAYCYFLVPHLDBLSGFAARRLAALIKGE

GGGMKIALHDLEIRGAGNILGTDQSGHIGTIGFNLYCKLLKKAVSALKKHTSPLLFNDDV

KIEFPYNSRIPDTYIETGSMRIEFYQKIGNAESSEELTAIQEMRORFGELPOEICWLFA

LAEIRLFALQHGISSIKGTANALYVOKCLSKSEQTKKTLDFALSPTPELLVKEVIESIER

GFLINAS

CPn_0892 1023673 1021046
ala5-Alanyl trna Synthetase
EFFFMLSNTIRSNFLKFYANRHITILPSSPVFPHNDESILFTNAGMNOFKDIFLNKEKVS
YSRATTSOKCIRACGKHNDLDNVGHTSRHLTFFEMFGNFSFGDYFKAEAIAFAWEVSLSV
FNFNPEGIYATVHENDDEAFALWEAYLPTDRIFRITDKDNFWSMANTGPCGYCSELLFDR
GPSFCNASPLDDTDGERFLEYMULVFMEFNRFTSEGSLLALPNKHVDTGAGLERLVSLIA
GTHTVFEADVLRELIAKTEOLSGKV/HPDDSGAFRVIADHVRSLSFAIADGLLFGNTER
GYVLEK ILRRSVNYGRGLGFRNFFLAEIVPSLADAMGEAYPELKNSLQIQK/LTLEES
FYKTLDRGVNLGGOVLKSSSSSSCIGGEDAFLKNTYGMPIDEISLLAKDYD/SVDMDTF
HKLEOFAKERSHKNVVQSGGTSESIYNELHLYTSEFTGYDHLSGDTFIEAIICKDHTVSSL
DEKOGJAIVLKVUPFYAEKKXQVGLGGEIFGGEJTFIVTHTTISKAGLIVHHKRISGSL
TVEAAVTAQVNRYRRRIALANNITACHLLIKALEITLGDHIRQAGSYVDDTKIPLDFTHPO
ALDFEOLL-TETLVNESIRENEFVDIRFERLYSDENNSSELKQFFGGKYJDVVF/YGAGHS
HELCYTTHAENTGDIGFFRITKEHAVAN ITRIEREAVTGHEARTVHQGGSVLEEIATLGH
HELCYTTHAENTGDIGFFRITKEHAVAN ITRIEREAVTGHEARTVHUGGSVLEEIATLGH
HELCYTTHAENTGDIGFFRITKEHAVAN ITRIEREAVTGHEARTVHUGGSVLEEIATLGH
HELCYTTHAENTGDIGFFRITKEHAVAN ITRIEREAVTGHEARTVHUGGSVLEEIATLGH
HELCYTHAENTGDIGFFRITKEHAVAN ITRIERATUTGHAENTVHUGGSVLEEIATLGH
HELCYTHAENTGDIGFFRITKEHAVAN ITRIERAVTGHEARTVHUGGSVLEEIATLGH
HELCYTHAENTGDIGFFRITKEHAVAN ITRIERAVTGHEARTVHUGGSVLEEIATLGH
HELCYTHAENTGDIGFFRITKEHAVAN ITRIERAVTGHEARTVHUGGSVLEEIATLGH
HERCYTHAENTGDIGFTETTENDAGLIFTGLITTELANTGHAENTUNGGGGAVELEIATLGH
HERCYTAGSTARALTATEVLNIFTLWWITGTGLITTELANTGHAAVALTRIERERATVHUGGGSVLEEIATLGH
HERCHARDSVARGARALTGHENDERLIFTENGERTATERATUTGHAAVALTRIERERATUTGHAENTUNGGGAVLEEIATLGH
HERCHARDSVARGARALTGHENDERLIFTENGERTATERATUTGHAATGHAATEVALTERE

Opn_9893 1025888 rktB_TransketoTase EFLAFCLG LOYSCOFY LEG INKELLA TILLGE TOLGALAY LLYGVURONF INKERF ZUGAGHGSALLYSCLHLAGEDVSLEDUGE PLOCAELAY LLYGVURONF INKERF ZUGAGHGSALLYSCLHLAGEDVSLEDUGE FROLHSKTPÖHPEYGETVÖVEATTÖPLGSALLINAVIMALSHKMLESRFINRTHEI FROKT YCLAGDOCFMEGYSHEVOSFAGSLALLNILWI YTOYNNIVALDGYLINELSUEDTKKRFEAYG WIDVYEI DGYDFTH HETFSSI KROGERSTULVIAHTI I CHGSPKETINAHGSPLOVEDTH ETKOFMHLEEEKF FVPPAVKNIFFAHK LØEDRKADEOWLDEVRIWSKOP PELHEEFVALTS HKLDKNLESLVOSVEMPDSI AGRASANKLI LOVLVOH I PYLLIGGSADLSSSOTHTANEKV HKLPKNLESLVOSVEMPDSI AGRASANKLI LYVLVOFREFOGT FLOYFSDYMRNATRLAALSKLP HTYDFSGRNIL TOLGALAGE FOR THE AGRADINELSUE TOLGALAGE FOR THE AGRADINELSKAD TOLGALAGE FOR THE AGRADINE THE AGRADINE THE AGRAD TOLGALAGE FOR THE AGRADINE THE AGRADINE THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR THE AGRAD TOLGALAGE FOR

CPn_0894 1036823 1025888
amn-AMP Nucleosidase
PRIDKNAKNILRRKHYKGEN/SKHTSESRIAQDMLERYSGSSVKQFCPYLLLTNFSYYIOT
FAKLHGVPVFEGSMFSAAJAPHLKTSILDFKLGSFGAALTIDLCSFLPDLKAALMLGMCG
GLBSHCVCVDYFVPVAS/RGEGTSDAYFPPEVPALANFVVQKATTEVLEDKKANYHIGIT
HTTNIRFWEFNKKFRKJLYETKAQSAEMECATLFAAGYRRNLPIGALLLISDLPLRKEGI
KTKSSGNFIFNTYTEDHILTGQEVIENLEKVMLKRAASDHKKDQQYRGLPHMEVGEADDT
MAGGSETSDSDY

CPn_0895 / 1026973 1027557
efp-Elongation Factor P
EIDCFMVRVST9EFRVGLRIEIDGOPYLILONDFVKPGKGQAFNRIKVKNFLTGRVIERT
YKSGESVETADIVERSMRLLYTDQEGATFMDDETFEQEVVFWEKLENIROMLLEDTIVTL
VLYNGDVVAVPPIFMELSIAETAPGVRGDTASGRVLKPAVTNTGAKIMVPIFIDEGELV
KVDTRTGSYZSRVSK

CPn_0896 1027574 1027822
CT753 hypothetical protein
EKYFFFFFFFFFRMEAKKIKELSKEAQLLKKLREKSRVLDEKNKRKAWVAKLVAMPESIREIE
KEERVETPQLFQAIAEKILEEGV

CPn_0897 1028794 1027853
(phosphohydrolase)
nfildsntvdoknksnprpmoekprhvhriihisdvhfhvlpvnpvhcfnkrlkgllrkv
nfildsntvvdoknksnprpmoekprhvhriihisdvhfhvlpvnpvhcfnkrlkgllrkv
ffilhfoattigorfpktvrslgadsvcitgdfsltamdgefllakhfvetlakhssvyl
legnhdvytlkslagoffythfphodlognkvsfhkitdhwhlilldcsclmgwfsangv
hlagisaietfllslspeenviiahhypelssoppshdlinnthlonvlkkypkvalyl
hghehoaavyncadtspsyilnsgsislptnsrfhvidlypekyqvhtmilknlldfdap
leianeatwdcokl

CPn_0898 1030511 1028904
Mitochondrial HSP60 Chaperonin Homolog
TKKRIGSVKILRLLGVCMSEDEKLSNYMADKKLFSGIDKLFQIVKGSYGPKQSLSPTSFF
KERGFYAISQTELSMSYENLGVDFAKAMYNKIHKEHSDGATTGLILLHAILQESYAALEK
GISTHKLIASLKLQGEKLOBALQQOSWPIKDALKVRNIIFSSLHMETIADHFYNAFSVVG
PEGLISITKERENDKTSMDVFQGFKIPAGYASTYFVSDTASRLTRIAHPLILITDRRISH
HISLLPLLQEISEONQHLIIFCEDIDPDVLATLVVNKLQGLLQVTVVTIPQLSTTNQELA
EDIALFTCTHICPCQEASHVLAPEMYTLGSCLSIEISESQTTLIGGLHIPEVLTLKTRQL
AEEIRTTSCLETKKRLIKSTRRLQSSVAILPTDEDNEPLYTLALKIMESALSRGYVPGGG
VALFYASLTLGTPKDDADENSIAISLLQKACCAPLKLLATNADLDGDAVIAKLSSLGTTS
LGISVFSREIEDLIAGGILDSLATTSTILAQALDTAILVLSSKILILENQYEISTL

CPn_0899 1030848 1032215
muff-Muramoy1-DAP Ligase
MHRCCRQMYMRAMLLEDAVSIMLSDVSCPKCDKKITGFAIDSQQVQPGDLFFALPGNATD
GHQFLKHAATAGAVAAVVSHDYQGDSFGLELIRVDDTKSALQEAGSNQCNLFCQTTLVGIT
GSVCKTTTKEFSKTILSSIYKTHASPKSYNSQLTVPLSLLMAEGDEDVMILEMCVSEPGN
MQDLLRIVQPEIAVITHINDQHAMHFPQGIGEILKEKSYILQKSKLQLLPKDSPYYLDLR
SCSPTAEKFSFSFNDPLADFCYKAISGDSVVIQTEEBHYCLPIAFSYKPAYTNLLIAVAL
SWILEVPEEGVIRSLPELKLPPMRFEHSMRNGMQVINDAYNACPEAMIAALDALPLPSOK
GKIILIIGHMAELGRYSEEGHALVAEKAASRGDMIFFIGEKWIPVQSVLKSYSCEVSFFS
SAQDVKDILKQVARYGDVILLKGSRALALESLLACF

CPn_0900 1032208 1033281
mraY-Muramoyl-Pentapeptide Transferase
LVFNFLGASMIPLIEMFLKQSLFFSLALTGMTTLVLTVALGVPVMKWLKRKNYRDYIHKE
YCEKLEMLHKDKAEVPTGGSVLLFISLIASLLVWLDWGKFSTWFFIILLTCYAGLGWYDD
RIKIKRKQGHGLKAKHKFMVQIAIAAFTLIALPYIYGSTEPLWTLKIPFMEGMLSLFFWL
GKVFCLGLALVAIIGTSNAVNLTDGLDGLAAGTMFSFAALGFIFVALRSSTIPIADDVAYV
LAALVGAGLGFLWYMSPFAOLFMGTGSLLLGGLLGSCAVMLRAECILVVIGGVFVAEAG
SVILQVLSCRLRKKRLFLCSPLHHHYEYQGLPETKIVMRFWIFSFVCAGLGIAAVLWR

CPn_0901 1033239 1034537
murD-Muramoylalanine-Glutamate Ligase
FCMRRSRYSCCLMEIDMCQRILILLGTGITGKSVARFLYQOGHYLIGADNSLESLISVDHL
HDRLLMGASEFPENIDLVIRSPGIKPYHPWVEQAVSLKIPVVTDIQVALKTPEFORYPSF
GITGSNOKTTITTLFLTHLLINTLGIPAIAMGNIGLPILDMHQDPSVRVVEISSFQLATQEE
HIPALSGSVFLNFSRNHLDYHRNLDAYFDAKLRIQKCLRQDKTFWWEEGSLGNSYQIYS
EEIEEILDKGDALKPIYLHDRDNYCAAYALANEVGWVSPEGFLKAIRTFEKPAHRLEYLG
KKDCVHYINDSKATTVTAVEKALMAVGKDVTVILGGKDKGGDFPALASVLSQTTKHVIAM
GECRQTIADALSEKIPLTLSKDLQEAVSIAQTIAQEDTVLLSPGCASFDQFQSFKERGA
YFKLLIREMOAVR

CPn_0402 1034507 1035241
nlpD-Michailduse (invasin repeat tamity)
AVDQRINASSEVNNNRRDMYTTAVVNNALLLVALEVTSKRIGVKDYDEGFPNFASSKVTQA
VVJSEKVIERPVVAEVPSRFIAKETLAAGFIESKEVIVTTEPVPVVJSETFEVPTVAVPEQ
EVRETVKEEQADYATVVVKKGDFLERIAFANHTYTVAKLMO INLDLITTOLK RGQYIKVTTS
ODVSNEKTPOTOTANI-ENYYIVZEGDCFWTLALRNHIRLDDLLKMNDLLEYKARRLKFGD
OLRIP

1035> Y 5

CPO_0003 10.05245 163-0417
EESW CELL DIVISION PROCEIN FESW
CKQCLCMKWFVISCELERIESIGEINVERFOSSEVEDROLECCTHKALTFOVTYLLLOLGV
CKQCLCMKWFVISCELERIESIGEINVERFOSSEVEDROLECCTHKALTFOVTYLLIGIGIV
ACLLYMMEMRIFICK CHVILLSDAALALTOFF FEGLOLICRINARIAWESFOOLT LQUEEFVK
YLOFF FULLYFETPESSEVEDY OKOLKMENTATE FEFTELLIALEFENGSAAVIGASIGEIVFFIM
TOVRERVWELFELGVULAGGALAVAMFOVRYRENVYLHPELDIKGRUIGFOAKIAACSG
KELDIKGROGSEGKEPVEPEANNOVIAALTAEEFGFICMEVELLLUMCFOONVALALKAS

SSLIANMCGVTLLLKV

1036320 1037396

CPn_0904 1036320 1037396
murG-Peptidoqlycan Transferase
RYMMKKIRKVALAVOSGOHIVPALSVKEAFSREGIDVLLLGKGLKNHPSLQQGISYREI
RYMMKKIRKVALAVOSGOHIVPALSVKEAFSREGIDVLLLGKGLKNHPSLQQGISYREI
RSGCPTVLNPIKIMGRISLGSGYLKARKELKIFDPDLVIGFGSYHSLPVLLAGLSHKIP
DEGULVOSKOMIGLEGRYARGIGVNESPVTKHFRCPAEEVFLEKRSFSLGSPMMKRCT
DETTETLGVVCKSSA WAD LATEGYPALVELVBSTEAMSTALEVETENDVMFVAUTVERFED
DEGUVEFPERGULVLLAADEVIGHASMTILEEBILDARVOSTLIPPTRAYSIGSETIAKEEV
DEGUVEFPERGULVLLAADEVIGHASMTILEEBILDARVOSTLIPPTRAYSIGSETIAKEEV
DEGUVEFPERGULVLLAADEVIGHASMTILEEBILDARVOSTLIPPTRAYSIGSETIAKEEV DVLEGGTMILEKELTEKLLVEKVTFALDSHNREKORNSLAAYSQQRSTKTFHAFICECL

CPn_0905 1037400 1039835
murCaddla-Muramate-Ala Ligase & D-Ala-D-Alam Ligase
MurCaddla-Muramate-Ala Ligase & D-Ala-D-Alam Ligase
MYMMKCTPOYHFIGIGGIGMSALAHILLDRGYEVSGSDLYESYTIESLKAKGARCFSGHD
MYMMKCTPOYHFIGIGGIGMSALAHILLDRGYEVSGSSKIFVAEADESDGSLKHYTPRAVVIT
TSSLIRAIFOEAQKDFSYAIGGLAANCLMGYSGSSKIFVAEADESDGSLKHYTPRAVVIT
NIDNEHLNNYAGNLDNLVOVIODFSRKVTDLNKVYYNGDCPILKGNVGGISYGSPECQL
HIVSYNQKAMQSHFSTTFLGQEYQDIEIALPGONNANAAAAGGVALTTGIDINIIRKAL
KKFSGYHRALERKNISESFLFLEDYAHHPVEVAHTLRSVRDAVGLRRVIAIFQPHRFSRL
EECLQTFPRAFOEADEVILIDVYSAGESPRESIILSDLAEQIRKSSYVHCCYVPHGDIVD
YLRNYIRIHDVCVSLGAGNIYTIGEALKDFNPKKLSIGLVCGGKSCEHDISLLSAQHVSK
YLSPEYDVSYFIINRQGLWRTGKDFPHLIEETQGDSPLSSEIASALAKVDCLFPVLHGP
FGEDGTIQGFFEILGKPYAGPSLSLAATAMDKLLTKRIASAVGVPVVPVOPLNLCFWKRN
FGELGTOLIETTSFFRUVKTAHLGSSIGIFLVRDKEELQEKISEAFLYDTDVFVEESRLG
SREIEVSCIGHSSSWYCMAGPNERCGASGFIDYOEKYGFDGIDCAKISTDLQLSQESLDC
VRELAERVYRAMQGKGSARIDFFLDEEGNYWLSEVNPIPGMTAASPFLQAFVHAGMTQEQ
IVDHFIIDALHKFDKQOTIEQAFTKEQDLVKR

1040514 1039915 CPH_0906
CTT63 hypothetical protein
kWGSEVLELVNDSQLSREASAFRLDIDFFILNIYPFFRNFKNIELCFFLSISQFNLDFME
EFVAYIVKNLVTNPEAVEIRSIEDEDNESIKLEIRVAAEDIGKIIGRRGNTIHALRTILR
RVCSRLKKKVQIDLVQFENSTDVIADQDYICDNDSSNSTEDTFGESDTCCSGHCHYDEDL NOEEQEEGNMHHSCECSNHH

CPn_0907 1040816 1040445

*CutA Periplasmic Divalent Cation Tolerance Protein CutA (C-Type Cytochrome Biogenesis Protein)

FAFSKFLIIKSSMTAVLILTSFPSEESARSLARHLITERLASCVHVFPKGTSTYLWEGKL
CESEEHHIQIKSIDIRFSEICLAIQEFSGYEVPEVLLFPIENGDPRYLNWLTILSYPEKP
DI SP:

CPDE 1908 1041607 1040780
CPDE 1908 1041607 1040780
CTTG4 hypothetical protein
ILAT FMILIKNBELMIRRFFKTLFPPGPQYSLCYASILIVLSSLVCVPTFCWLFLPELS
LSKTMPSPIRNLFLVSSTLSKVPPTAIAEHLRLSADAPTYLHEFSIKEAESSLHALGIFS
SLYZEKSPDMKGTTIFYTLOTPIAVYGRRSNTLCNLEGSCFLGQPYFPSLNLPQIFFSQE
DLGGGKLPKEKMLFTKILLKELAMESPKIIDLSLSDAYPGEIIVTLSSGSLLRLPIKTLD
RALBLYKHMKKSPVIESEKQYVYDLRFPNFLLLKAL

CPA-0909 1041592 1041966
rsby-Sigma Factor Regulator
IISTIFTRFLLERLLMNLSAKEYGDIIVIYLQGSLDAVSVPSVQEYLEQFIQKKHLKIAL
NPTDYSVISSAGIRLLLSNFKLVQSLGGKMCLCCVKESVTEVMRIAGLDQLILLCQSEQE
CLSKL

CPD 0910 1041970 1043004
miak+trna Pyrophosphate Transferase
FLYMLPFEFERNTTSSPECDVCLDPOKLFVKLFKRTIVLLSGPTGSGKTDVSLALAPMID
GETVSVDSMQVYQCMDIGTAKVSLKARQEIPHHLIDIRHVQEPFNVVDFYYEAIQACONIJ
LSRKVPILVGGSGFYFHAFLSGPPKGPAADPQIREQLEAIAEEHGVSALYEDLLLKDPE
YAÖTTTKNDKNKIIRGLEIIQLTGKKVSDHEMDIVPKASREYCCRAWFLSPETEFLKNAI
QMRCEAMLQEGLLEEVRGLLNQGIRENPSAFKAIGYREWIEFLDNGEKLEEYEETKRAFV
SNSWHYTKKOKTWFKRYSIFRELPTLGLSSDAIAQKIAKDYLLYS

CP# 0911 1044079 1042985 CHICAGO CONTROL ON THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF T SLIEAIFNVNYFMICKRISFEEGLELFVSSPIERLQERADAIRKERY PSNEVYVLDAN PNYTNICKIDCTFCAFYRKPKSPDAYLLSFDEVRSELLQRYVSSGYKTULLOGG/VVLDAN PNYTNICKIDCTFCAFYRKPKSPDAYLLSFDEVRSELQRYVSSGYKTULLOGG/HPGLGI DYLLEELVRITVQEFPSIHPHFFSAVEIEHACRVSGISIEQGLQRLMDAGQRYLFQGAEI LSERVRKIISPKKMQPGGWINLHKLAHLMGFRTTATMFGHVENPEDILIHLYOTLRDAQD SCPGYYSFILMGKRIFTMSVKPGTALRRNVPQQASIETYYRILALGRIFLDNFDHVAASWFGEGKS LGAKALHYGADDFGGVILDESVHKATGWSIQSSEEEICNIIRSEGFIPVENNTFYQHISC

1044120 1045760 CPH_0912 1044120 1045760
CTT68 hypothetical protein
VVIMDNSDNSFHTLETEQGSFLNDELAVEEVASTESTEISDATLCSAEKKVAFILNKMRE
ALMOSSQSGDLRLFWDLRKQCDFLFNEIEDTAKRADHWRCYIELTKEGRHLKGLQDEEGS
FVVQGIDLAITCLEKDILKFQEGTEDKIFKDREDNFLESQALDKAQAFYKQHHTSLLWLS
SFSSKIIDLRKELINVOMRNRLKSKFFORLSNLGNQVFPKRKEJIEKVSOFFAEDVDAFV
AKYPIGSBKETLKKTVFFLRKEIKKNCHAAKRLFVSSHVFAEFTRIKLSKKUPOLKGMEKE
IRQEQGRLRVVSAENSKEVRQMLAEVSSLLIEGNDLSKVRKJLEGISKKIRALDLTHDDV
ISLKKENQOLFDOLREKQDAEHSYQEQLAKDKQVKKEARSFLAERITTFSKTSEGNIT
SESREEWOTLKELLGKMSFLPPEKISLDNQLNLALQTIVMFFEEQLLSSPDSREKLVMM
RQVLKORRERQELKDKLEQDKKLLGSSGLDFDRAMQYSALVEEDKRALEELDASILELK
DOLOOLI

CPn_0913 1045709 1045945
No cobust homolog present in Genebant/EMBL as of 11/7/98
RESCRYPRIEATDSALAMRENCIYAFDLOGTLERGGSWSFYCYGLLAGLESYKTLPPCI VREEKEKEEGIEHESIIR

Ctn_0914 1045999 1046399
No cobabt homotog present in dendernak/EMBL as of 1177798
VERWOLESHYD IVTRILLSSYCODDLYEVALDEVSTLITSSDEVARVLEKLEEAFADTTYO
VILESSIEDE IVINITAQUIATISSWYASSYRBOSAEVITYKKCLTYDKKAQTLSYTKKINQ
ARGITTSDHILDI.FFLMIGEEKTVVRPQSYLKKMAYKYYWNIV

1046401 3046817 CPn 0915

OO LOOPLY

YBBB LOJAP SUPERLAM POLICY FOLKKSTUSWI LEYPKAGE FOLKVAAK IDDKKONNLVVLOVRTISEFTDYFV FVEGSVNVHVKALAIT IVEELKKOKVSFLIVEG IDDKMVVI DYGFIVVHVFVSEIRGKY RLEELWKDGFIVTSKLLAS

CPn_0916 1046813 1048984

£abf-Acyl Carrier Protein Synthase

LLNVRVYMSKRVVVTGGTVVSCLGNEVDFFYDNLLAGVSGVRPITSFPCEDYATRAG

LLNVRVYMSKRVVVTGGTVVSCLGNEVDFFYDNLLAGVSGVRPITSFPCEDYATRAG

SCHEENBESCHE ABESTA ENTARC VARIA LAMERICHER STELLYGGENE

SCHEENSENDER STELLY LYBMAFAL LAMER LAMER STELLYGGENE

DAAYGHLVSGADMIICCGTEAAVNRICHEGFIANKALSENDAPCJASRFWORDROGFV

DAAYGHLVGESALRRDAPIFAEMLGSYVTCDAFHITAPRDDGEGITACVLGALNSA

GIPKERVNYVNAHGTSTPLGDLSEVLÁVKKAFGSHVRNLRMSTKSLIGHCLGAAGGVEA

VVALQAILTGKLHPTINLDNPIAEIEDFDVVANKAQDWDIDVAMSNSFGFGGHNSTILFS 1046813 1048984

1048064 1048539 CPn_0917 1048054 1048539
hydrolase/phosphatase nomolog
frddilevctlvmmktkyeysgvlfbikffctpdkntlkacfichtrgkhwgfpkghsed
keopoeaaerelveetglsvynffpkvliegysfnneegyfvrkevtyflaevrgdihad
pmeicdsgwlsloeglrllsfpelrdltveadkfinnylfss CPn_0917

049232 1048579 CPn_0918 CPT_0918

Dpa-Inorganic Pyrphosphatase

ELLMSKKPLYVAHPWH9PTLTODNYESLCCYJEITPYDSVKFELDKATGLLKVDRPQKFS

ELLMSKKPLYVAHPWH9PTLTODNYESLCCYJEITPYDSVKFELDKATGLLKVDRPQKFS

NFCPCLYGLLPQTYCCTASGNYSGEQTRREGIQGKDPLDVCVLTEKNIHHGNILLQARP

IGGLRIIDSGEADDKJIAVLEDDLVFAEIEDISDCPGTVLDMIQHYFLTYKATPNHLIKG

SPAKIEIVGIYGKKFAQKVIQLAHEDYLSYIGDTAEVN

1049375 1050430 CPn 0919 CPn_0919 1049375 1050430

1dh-Leucine Dehydrogenase
FMYSLNFKEIKIDDYERVIEVTCSKVRLHAIIAIHQTAVGPALGGVRASLYSSFEDACT
PMXYSLNFKEIKIDDYERVIEVTCSKVRLHAIIAIHQTAVGPALGGVRASLYSSFEDACT
DALRLARGNTYKAIISNTGTGGGKSVIILPQDAPSLTEDMLRAFGQAVNALEGTYICAED
LGVSINDISYVAEETPYVCGIADVSGDPSIYTAHGGFLCIKETAKYLMGSSSLRGKKIAI
GGIGSVGRELLQSLFFEGAELVVADVLERAVQDARRLYGATIVPTEEIHALECDIFSPCA
RGNVIRKDILADLNCKAIVGVANNQLEDSSAGMMLHERGILYGPDYLVNAGGLLNVAAAI
EGRVYAPREVLLKVEELPIVLSKLYNQSKTTGKDLVALSDSFVEDKLLAYTS

CPL_0920 1051423 1050431

Cys0_Sulfite Synthesis/biphosphate phosphatase
ILEENSHHSELPNYONIVESVYTEITTQLLNYRSEHRLVPFWEKSDGSFITAADYGSQYY
LKOLAKAFPNIPFIGEETLYPDODNEKIPEILKFTILLTSSYRDDLISTLVPPPSPTS
LFYLVDPIGGTAGFIRHRAFAVAISLIYYEYRFILAWACPAYNOTFKLYSAAKGHGLSIV
HSONLDRRFYYADRKOTKOFCEASLAALNOOHHATRKLSLGLPNTPSPRRVESQYKYALV
AEGAVDFFIRYPFIDSPARAWDHVPGAFLVEEAGGRVTDALGAPLEYRKESLVLNNHAVI
LASGDQETHETTLAALQNQLNVVPTDKLIAL

1051526 1052293 CPH_U921
snGlycerol-3-P_ACyltransferase
GELMLIKLWRATYEGMYTFLVGALLKLRYRMQVEGMDTLNINPKQGCLFLANHVAEVDPI
ILEYLFWSRFHVRPMAVEYLFHSRVVOWFLBSVRSIPIPOLVPGKESKRSLERMNVCYEE
ASRALNRGESLLLYPSGRLSRTGKEEIVNQYSAYVLLHRVWECNVVLVRVSGLMGSAFSR
YKONSTPKLGPAFKEAFRALLRRGIFFMPKRFVKITLCQVDHLFLKQFPTKQDLNTFLAS

CPn_0922 1052266 1053927

aas-Acylglycerophosphoethanolamine Acyltransferase
OFAHRSSLRITRKLRRHHDORNRGHNNINLRLRPGSTLLEAFLILCSEHEEGIACFDEHL
GSLSYRELRNAIIAVAIKVSKFSEDRVGVMMPASIGAFIAYFGILLAGKTPVVMMMSOGL
RELRACTKTVEVRRVLTSQOFIKHLTEVOGFVEYPFDLMYMEDVRKRLSWMEKCRIGLYS
KCSVPMLIRIFGVSGVESDDTAVILFTSGTEKLPKAVPLTHKNLMENQEACLKFFDPNTQ
DVMLAFLPPFHAYGFNSCGLFPLLMGVHVVFASNPLNPKKLVEFIDDKKVTFFGSTEVFF
DYILKTAKKONSCLESLRLVVIGGDALKDTLYEETKKLQPQIALYQGYGATECSPVIST
KESPRKSECVGMPIEGEMDVLIISKETHIPVSSGEQGLIVVRGNSVFSGYLGNHEHQSFV
SLGGDQWYLTGDLGHIGPSGDLFLEGRLSRFVKIGGEMVSLEALESILHEHFTENQMEDA
GSLVVCGIPGBKVRLCLETTLATTIHEVNDILKSAETSSIVKISYVHQVESIPILGIGKP
DVYSINALAVSLFG 1052266 1053927 DYVSLNALAVSLFG

1053966 1055093 CPn_0923 1053966 1055093
bioFl-Oxononanoate Synthase_1
VKESFLTTSDVIDFVTMDFLGFARSPTIYCEVSKRFQIHCQQFPHEKLGIRGSRLMVGP
SSVIDDLESKIASYHGAPNAFIVNSGYMANLGLCHHVSRSTDVLLMDEEVHMSVVHSLSA
ISGOHHTFHHNNLEHLESLLQCYRISSKGRIFIFVSSVYSFRGTLAPLEQIIALSKKYHA
HLIVDEAHAMGIFGDDCKGLCHALGYENFYAVLVTYGKALGTMGASLLTSSEVKYDLMQN
SPPLRYSTSLSPHTLISIGTAYDFLASEGEIARKQVFKLKEHFHECFDSHAPGCVQPIFL
PHTCLEEAISVLETTGIHVGVVAFAKHPFLRVNLHAYNTVDEVNLLAQVMKPYLEKSSHR VHINHEFHLWRELCCH

CPn_0924 1057301 1055028
priA-Primosomal Protein N'
KRFTAKTKSMGYIESSTFRLYAEVIVGSNINKVLDYGVPENLEHITKGTAVTISLRGGKK
KRFTAKTKSMGYIESSTFRLYAEVIVGSNINKVLDYGVPENLEHITKGTAVTISLRGGKK
KRFTAKTKSMGYIESSTFRLYAEVIVGSNINKVLDYGVPENLEHITKGTAVTISLRGGKK
VGVIYOIKTTTOCKKILPILGLSDSEIVLPODLLDLLEWISGYYFAPLGKTLKLFLPAIS
SN/LOPKOHYRVVLKGSKAKTKETLAKLEVCHPSGGAVLKILLQHASPPGLSSLMETAKV
SQSPIHSLEKIGILDIVDAQLELQEDLLTFFPPAPKDLHPEQQSAIDKIFSSLKTSGFH
THLLFGITGSCKTEIYLRATSEALKOGKSTILLVPEIALTVOTVSLFKARFGKDVOVLHH
KLSOSDKSRTWRQASEGSLRILIGPRSALFCPMKNLGLIIVDEHDPAYKOTESPCYHA
EDVAMMCKLAHATVYLGSATPSLESYTMALSGKYVLSRLSSRAAAHPAKISLINNDLE
PEKKIKTK LLFGSPVLKKIAERLEVGEOVLIFFNRRGYHTNWSCTVCKHTLKCPHCDMVLT
FHKYAFFLLCHLCNSSPKDLINGCPKCLGTMTLOYRGSGTEKIEKILQJFFDIRTIRID
SDTTKFKGSHETLLRGFATNIKADVILGTOMIAKGNFSAVTLAVILLNGGIYIPDFRTIRID
SDVTPOLTTOVAMRSGRBILLYBEILIGGFLPDHPTIHSAMPQDYSAFYGGEITGRELCEYP
PFIRLIPGIFMCKCFKGTWEEAHRVHNILKEQLESTNPLMF/TPCGHFKIKDTFRYQFLI
VGAYVIPVNKKHHALMLAKLDPKVKFMIDVDFMTTFF

CPO_0925 TOS7915 1057226

CT779 hypother ical profess
RHMLEMENSONEHDTLCOLLDRYGEDLYCTLAGLENVTLPHTATGAGYGSTPEKAVEVPN
REPORTPPPTMILGSEKTRIGHANGCVPLHFDLGONATEKEKYPALYCGSTPARKTPGST
FYVERINEEVLEFNRLAKTLTGGLEFPTRIGTLHAKTNIFYINPREFLALADLMYERYKTD
TTDYHGGLTONGCTFLPLYGGLEYCKDGOLKRILMATLINRLPFAYTPYGS 1057915 1057226

1058000 1058557 CPn_0026 1058000 1058 Thioredoxin Disulfide Isomerase THEOTOTY:TREFKDSDMKFWLOGCAFVOCLLLTLPCCA. KASGENLQOTRPIAAANL OWESYABLEHSKOOMKPTCLFFTGSDMCMMCIKNODQILQSSEFKHFAGVHLHMVEVDF PQNHQPEEQRQKNQELKAQYKVTGFPELVFIDABGKQLARMGFEPGGGAAYVSKVKSAL

1050918 1058670 CDn_0007 427

CPD_0027 437 1050918 105870

**HERE TO REFERENCE TO THE MEMBERS OF THE TENTH OF THE MEMBERS OF THE TENTH OF THE MEMBERS OF THE TENTH OF THE MEMBERS OF THE TENTH OF THE MEMBERS OF THE MEM VDRLAAVISNYLDSENRKSQQPD

CPn_0928 1061035 1059884

*CHLPS 43 kDa protein homolog_3

RKDFAFTLINLSNRSDILSGIFSNPHPVSYFSSTHAKQLSDFSKKHPILTKIVTIIVKI

FKLLIGLIIPPLGIYWLCQLVCSLALFPRSSMLYSVLKTCFKKYRLEGEIQDYFVKNLDP

SFKDPAVSESKRITIQQDHITIOTLAIHFSTARPKRWLLISLGSGDFLEMIGLKDSFL

SWKELAKLLGANILIYMPGVKSSTGKLMLENLATAHNICAKYLQDKIQGGANEIITYG

YSLGGVVQSAALQKNPFTNSETSWVAVKDRAPHSLPAAANSFFGPIGKLIAVLARWKNDA

EKNSRELFCPEILVYSADRFRPSEVGDDTALLPEFTLAHAIKRTPFARSKKFIGEVNLLH

SSPLKHPTIQKLAEAILESLSRKN

CPn_0929 1062301 1061186

*CHLPS 43 kDa protein homolog_4

KFMAPIHGSNAPVEDILHSHPSPQATYFSSTRAQKLHEFKDRHPVLTRIASVIIKIFKV
LIGLIILPLGIYHLCOTLCTNSILESKNILKIFKQPNTKTLKTNYLHALQDYSSKNRVA
SMRRVPILQDNVLIDTLEICLSQAPTNRWALISLGSDCSLEEIACKEIFDSWQRFAKLIG
ANILVYNYPGVMSSTGSSSKULDASAHNICTRYLKDKEQGPGAKEITYGYSLGGLIQAE
ALROQKIVANDDTTWIAVKDRCPLFISPEGFHSCRRIGKLVARLFGWGTKAVERSQDLPC
LEIFLYPTDSLRRSTVRQNKLLAPELTLAHAIKNSPYVQNKEFIEVRLSSDIDPIDSKTR VALATPILKKLS

CPn_0930 1062851 1063330
No robust homolog present in Genebank/EMBL as of 11/7/98
NKMSELAPCSTGLOMYPHTOVHHALDTRRVILTIAACLSLIAGIVLVGLGAAAILPSLFG
VIGMILILFSSTALTYLYKKTREVDQIALEPLPEMISKDQSIIDFVKTRDYASLEKKAT FAYTHTHYYDGSMVFYREIPRFMLGSYLALRKDMDRQALF

CPGT0931 1064078 1065718

1ys5-Lysyl trna synthetase
1DFgYLGWKSDIYTNILEERMTARAEYLDHEDFLYRSHKLQELSELGVVLYPYEFPGVFS
CEOLKKTFASGELGNSEAMSRSTPRVRFAGRLVLFRAMGKNAFGOLIDHNOTIQVMFNR
EFTSYHCLSEDAEITPIKFIEKKLDLGDILGIDGYLFFHSGELTVLVETVTLLCKSLLS
LPPKHAGLSDKEVRYRKRMLDLISSREVSDTFVKRSYIIKLIRNYMDAHGFLEVETPILQ
NIYGGAEAKPFTTTMEALHSEMFLRTSLEIALKKLLVGGAPRIYELGKVFRNBGIDRTHN
PEFFMIEAYAAYMDYKEVMVFVENLVHLVRAVNHDNTSLVYSYKKHGPGEVDFKAFMJK
MTMRESIATVAGIDVDMSDQKLKEILKKKTTFPETAFATASRGMLIAALFDELVSDMLI
APBHITDHPVETTPLCKTLRSGDTAFVERFESFCLGKELCNAYSELNDPIRQRELLEOH
TKKEELPDSECHPIDEEFLEALCQGMPPAGGFGIGVDRLVMILTNAASIRDVLYFFVMR
FDAEKTN

CPT=0932 1067160 1065721

CYSS-CYSCEINY1 tRNA SYNTHELASE
VKSDTVMAFSHIEGLYFYNTASOKKELFFPNHTPVRLYTCGPTVYDYAHIGNFRTYVFEN
VKSDTVMAFSHIEGLYFYNTASOKKELFFPNHTPVRLYTCGPTVYDYAHIGNFRTYVFEN
VKSDTVMKENSHVANNTTDVEDKTIACASKKNIPLQEYTQPYTEAFFEDLDTLNIARA
DFYPHATHYIPOMIOAITKLLEGGIAYIGGDASVYFSLNFFPNYGKLSHLDLSSLRCCZR
ISADEYDKENPSDFVLWKAYNPERDGVIYWESPFGKGRPGWHLECSIMABELLGDSLDIH
AGGVTNIFPHHENEIAQSEALSGKPFARYWLHSEHLLIDGKKMSKSLGNFLTLRDLHQE
FTGGEVRYMLLQSHYRTOLNFTEEALLACRHALRRLKGFVSRLEGVDLPGESPLPRILDS
SSOPEAFSRALANDLNVSTGFASLFDFVHEINTLIDGGHFSKADSLYILDTLKKVTVL
GVLPHTTSVCIPETVMQLVAEREEARKTKNWAMADTLRDEILAAGFLVEDSKSGFKVKPL

CPn_0933 1067532 1068578

predicted disulfide bond isomerase

PVILLONIRRCSLKQLKVLATLLLSLSLPTLEAAENRDSDSIVWHLDYQEALOKSKEAEL

PLUVIFSGSDMOPGMKIRKEVLESPEFIKRVQGKFVCVEVEYLKHRPQVENIRQONLAL

KSKFKINELPCMILLSHEEREIYRIGSFONETGSNLGDSLCHIVESDSLLKRAFFMWTSL

SLSELDGRYYRLAEELSHKEFLKHALELGVRSDDYFFLSEKFRLLVEVGKYDSEECQRIKK

RLUNKDPKNEKQTHFTVALIEFOELAKRSRAGVRQDASQVIAPLESYISQTGQQDKDNLW

RVEMMIAQFYLDSDQWHHALQHAEVAFEAAPNEVRSHISRSLEYIRHO

1068948 1068526 cri_0334
rnpa-Ribonuclease P Protein Component
rpvHpLTLPKQSRVLKRKOFLYITRSGFCCRGSQATF/VPSRHPGTCRMGITVSKKFGK
AHERNSFKRVVREVFRHVRHQLPNCQIVVFPKGHKQRPVFSKLLQDFINQIPEGLHRLGK TKATTGGECTPKSEKCVTAPR

1069100 1068957 CPn_0935 1069100 1068957 r134-L34 Ribosomal Protein EDTVKRTYOPSKRKRRNSVGFRTRMATRNGRKLLNRPRRHG RHSLVDL

CPn_0936 1069330 1069470 rl36-L36 Ribosomal Protein YLMKVOSSVKADPSKGDKLVRRKGRLYVINKKDPNPYSRQAGPARKK

1069487 1069798 rol4-S14_Ribosomal Protein
VKRMAKKOSVAKEAKRRLVEANFKKRODLRKIVKZLOVOEEEKENARISLNKMKROTSP THEIMPOLISTORPROYERKFALSRICFROMASMOPTPOYTKASW

1070355 106984 CPC_0038 https://doi.org/10.00395 106984/ CPTABE https://doi.org/10.00395/files/file CELIDEEELAATTIKGKAA EEE IMAMKKOWKD MÕKAKDAS OLOOGYTKEVSK

1071195 DFLLENSEC

1071204 1073033 CPn_0940 DATTE LUMPHEFECTION OF AVENERABLY ARE EMPLATEMENTED OR ERIPFLMKKTAS IETIVVSNET FALLLENNLIKCHHPKYNVLLKDKTFFCLAISLSHSW PKVEAIRTKAITSSORQLIFGFYVSAEACHTLLEVISOMFPLRTCSDREFALRKRPCILY DMKRCLAPCVGYCTPEEYOGT DKAILFLKCKIEEVVKDLEKVIQKASDNLEFEQAANLY MKRCLAPCVGYCTPEEYOGT DKAILFLKCKIEEVVKDLEKVIQKASDNLEFEQAANLY MTLLIKOAMAKOQVEKPHFØNIDALGLYRHKORTILTLLTVRSCKLLGARHFSFF ENAO EDODLLSSFILOYVYSOPY PKEILTPLPLEFPTLSYVLNAESPPRLRSPKTGYGKELLD LAYRNAKAYAATTLPSSTUPYODFONILRMSQYPYRIECYDNAHMOGAHATGYVIVFBN GFDPKGYRTFSIDSEKTGYDLALLEEVLLRRFHSLTTALPOMIVVDGGKTHYNKTKKIIQ TLNLTGIOVYTIAKEKSHSRGLNKEKIFCETFPEGFSLPFTSNLLOFFOILADEAHRFAISKHRKKRGKALFEQEYIPGIGEVKRKRLLOKFKSWKQVMLSSQEELEAIPGLTKKDIAV LLARQKDFNKSD Toyer Excipations DAMPTOLICMERECECTE menne

CPn_0941 1075504 1073018

mutS-DNA Mismatch Repair

VMTEKKPTPMMEOMHOCKEKACDSVLLFRMGDFYEAFYDDAVLLSOHLELTLTORQGIPM
SGIPVSTVDTYVDRLIGKGFKVAVAEOFGEPAKEKESKKIGPMARDIQRFVTFOTLLSST
LLOEKFNNY IVAINRIGSLFGFACLDLSTGSFFIEECENTKELVDEICRLAPSEVLSONL
KYNKETALVMGLOOHLKUTLSTYADMAFEHKFASOKLTHTFOVASLDGFGLKGLVPAINA
AGGLLSYIODKLLLPTKHIAIPOTRGKOOKLLIDTASOVNLELLAPLNDPOGKNSLLRIM
DHTSTPMGGRLLRGILISPFYNPKEILVRODAVEFFIROVILRRIKTYLOVROIEERI
KTVTTGLAGPRDIGTLRDSFSAGAQIYEOLASATLPEFFIDKCSLDTKLASLIALLSKSL
NGDLPLRVSCRNIFVDEFHNDLKRLRHNGEHSQEWIWEYQERIKETGIKKLKIGFAQAL
GYYIEVSEFFAPOLPKOPITROSSLHAERFTTIELOOFGDDMSNISEKLOTLETOFFKOL
CSHILOLKTEILALSOSLADLDYIISLADLAHAQGYCRPHVDMSDTLCTYRGCHPVAKTL
VDTGKFIRNDTEMGSOTRMILLTGPNNAGKSTYIRQIALLVIMAOMGSYIPAKSAHU
UDKI SFRIGAGDNISKGKSTFMVEMAETANILHNATDRSLVILDEVGRGTSTYDGLAIAQ
AVVEXLLFTDKKKAKTLFATHYKELTTLEDHCPHVENFHAGVKDKAGQPVFLYEILKGHS
QKSYGIHVARLAGFPLCVVSRAQQILRQLEGPESITRPAQDKMQQLTLF 1075504 1073018 CPn_0941

1075955 1077754 CPÁ_0942 1075955 1077754

d/aG/prim-dna Primase
wcsitklrtamyteesldnlrhsidivdvlsehihlkrsgatykaccpphtektpsfivn
bagahyhefoggahgdaigflmohlgysfteailvlskkfovduvlopkdsgytppoglk
belrhinsbatfffryctyhlpearhaloytyhrofspdtidrhlgygeoslfloame
erkisoeolhtagffonkwflfarriifpyhdalchtigfsarkflensoggkyvntpet
pifkksrilfglnfsrriakekkvilvegoadclomidsgenctvaaoggtatfeehve
lsklovlkvfllfdsdeagnkaalrvgdlcotaomsvfvcklpoghdddsfluorgssgl
lalleosodyltflisemssypkfofrekallveeairoikhwgspilvyehlkolast
mwypedmylslanpovtaeponipilkokypkthphivmetdilirchlfosntkiltao
fyfypedfkhecorlfafmisyyekyrknvpfdeacovlsdsoilolltkrrinteald
tifvoslokmadrkwreockplslnoniodkkleiledyvolrkdrtiitlldpeselip CPA_0942

1077972 1078238 CT794.1 hypothetical protein FFMKSFKFLLPFLSVILCCGNLLSSPRSRAISVTESIGMSAVKTLVLSEKAHEFLEGIGY GVGASSILRDWQTQQWLEIESLLAQNEVM

1078503 1078997 CPn 0944 CPT_0944 10/8503 10/6997 No robust homolog present in Genebank/EMBL as of 11/7/98 IKIMMHRYFIPLLALLIFSPSLVRAELQPSENRKGGWPTQLSCAEGSQLFCKFEAAYNNA IEEGKPGILVFFSERPTPEFADLTNGSFSLSTPIAKGFNVVVLCPGLISPLDFFHKMDPV ILYMGSFLEMFPEVEAVSGPRLCYILIDEQGGAQCQAVLPLETKN

CPn_0945 1079001 1079660
CT795 hypothetical protein
SIFKNKILPSYFGHNFDQLRRHYMRIALSLLSLLMIFPIFGEESRPGSEDGNSNTQEIVG
SQDTQVCLYHSYEQGLQSAFIEGKFLVIVVLCNSGDDGQACTIGLSETCEEVLSVLSGSI
FSELANFVVLVPSGVNPLIYPPIEDPILAEIVKFKELFKDESFPTGLSIIVVGVTPEGPG
DIIEVSPVSLTVEEEETLPSEQTTEVESTSELQSEDPAIA

CPn_0946 1082816 1079745
glyO-GlyCyl tRNA Synchetase
GECOKKCYTLESFVSEHPLTLQSMIATILRFWSEQGCVIHOGYDLEVGAGTFNPATFLR
ALGEPYKAAYVEPSRRPQDGYGVHPNRLDNYHQLQVILKPVPENFLSLYTESLRAIGL
DLRDHDIRFIHDDWENPTIGAWGLGWEVWLHOMEITQCHYPCAIGSKEDTISGEITYGI
ERIAWYLQKKISIYDVULMDTLTYGQ ITQASEKAWSEYNFDYANTEMWFKHFEDFAEEAL
RTLKNGLSVPAYDFVIKASHAFNILDARGTISVTERTRYIARIRGLTRLVADSVVEWRAS
LNYPLLSLSSTSEPKETSESVVPMISSTEDLLLEIGSELPATFVPIGIQQLESLARQVL
TDHNIVYGLEVLGSPRRLALLVKNVAPEVVQKAFEKKOPMLTSLFSPDGDVSPQQQFF
ASGGVDISHYQDLSRHASLAIRTVNGSEYLFLLHBEIRLRTADILMQELPLLIQRMKFRK
FMYWDNSGVEYARPIRWUALYGEHILPITLGTITASRNSFGHRQLDPKKISISSPQDYV
ETLRQACVVVSQKERRMIIEQGLRAHSSDTISAIPLPRLIEEATFLSHPFVSCGOFSEQ
FCALDKELLIAEMVHNGKYFTHETSSGAISNFIVVCDNSPNDTIIEGNEKALTPRLTD
GEFLEKODLOTPLTTFIEKLKSVTYFEALGSLYDKVERLKAHQRVFSTFSSLAASEDLDI
AIQYCKADLVSAVVNEFPELQGIMGEYYLKHANLFTASAVAVGEHLRHITMQKLSTIGT
LLSLLDRLDNLLACFILGLKFTSSHDFYALRRQSLEVTLVCASRLPIDLASLLDRLDH
FPSTIEEKVWDKSKTIHEILEFIMGRLKTFMGSLEFRKDEIAAVLIDSATKNPIETLDTA
EALQLLKEEHTEKLAVITTTHNRLKKLLSSLKLSMTSSPIEVLGDRESNFKVULDAFRGF
PKETCAHAFLEYFLSLADLSNDIQDFLHTVHIANDDGAIRNLRISLLLTAMDKFSLCHWE
SVAV 1082816 1079745

1083433 1084059 CPI_0947 TOSA(3) LOSA(3) CONDUCTED SO CELEVIDADES OF CONTROL C

1085483 1084047 BSBAARETOKOOGRATGARARGITATA LALTOGO LETEGABAAAGENINAAREGRAVAYYYY ATOV OLAGONOGRATGARARGITARA BITGRAGGROOATGE BSBAARETOKOUDAEAFTBAA BITGRAGGROOATGE BSBAARETOKOUDAEAFTBAA BITGRAGGROOATGE YLDEADPAD IVHLHDWHVGLLAGLLKNPLNPVHSK IVE DPHLSHYOLFRDFOTSVLMKGALYCSDY ITTVSLTYV SG I ING IDEDVWNPKTDPALAVQYDASLLSEPDVLFTK SDYELHDA ILARNSVE RAVLYEKLGISSDYFPLI CVIGRIVEEKGPEFMKEITLIHAMENSYAFILIGTSONEVILMEFRNLQDCLASSPNIRLI LDFNDPLARLTYAAADMICIPSHREACGLTQLIAMRYGTVPLVRKTGGLADTVIPGVNGF TFFDTNNFNEFRAMLSNAVTTYRQEPDVWLNLIESGMLRASGLDAMAKHYVNLYQSLLS

1085887 1086483

CPn_0040

__CPn_0040

__CPn_00 AGIETITPLKEVAITVSRR

1086470 1087027 CPn_0950

pth-Peptidyl trna Hydrolase
PSLEDMMAKLIVAIGNPRHGYANTRINAGFLLADRLVEELQGPPFKPLSKCHALMTLVES
SSGPLVFIKPTTFVNLSGKAVVLAKKYFNVALSHILVLADDVNRSFGKLRLCFNGGSGGH
NGLKSITASLGSNEYWQLRFGVGRPLEBGVELSNFVLGKFSEEENLQLGSIFVEASTLFT

CPn_0951 1087113 1087457
rs6-S6 Ribosomal Protein
EFLMGKKENQLYBGAYVFSVTLSEEARRKALDKVISGITNYGGEIHKIHDQGRKKLAYTI
RGAREGYYYFIYFSVSPGAITELWKEYHLNEDLLRFMTLRADSVKEVLEFASLPE

CPn_0952 1087469 1087723
rs18-s18 Ribosomal Protein
GEMMNKPVHNNEHRRKRFNKKCPFVSAGWKTIDYKDVETLKKFITERGKVLPRRITGVSS
RFQGVLSQAIKRARHLGLLPFVGED

1087727 1088248

CPI_0953
T19-L9 Ribosomal Protein
FKGRMKQOLLLLEDVDGLGRSGDLITARPGYVRNYLIPKKKAVIAGAGTLRLQAKLKEQ
RLIQAADKADSERIAQALKDIVLEFQVRVDPDNNMYGSVTIADIIAEAAKKNIFLVRKN
FPHAHYAIKNLCKKNIPLKLKEEVTATLLVEVTSDNEYVTVLAQGKQTEENQEG

CPn_0954 1088259 1088708
ychB-Predicted Kinase
GRKVCYKDIMQYFSPAKINLFLKIWGKRFDNFHELTTLYQAIDFGDTLSLKNSMKDSLSS
NAMELLSPSNLIWKSLEIFRRETOIHQPVSWHLNKSIPLQSGLGGGSSNAATALYALNEH
FÖRHIPITTLQLWAREIGSDVPFFFLQEQH

CP 0955 1088612 1089175

CFTEU993 (fizme-shift with 0954) RAFPNPYSYNNIATLGSRNRKRGSFFFSSGTALGKGRGEHLFSIKKLNHKHKYVLYLDHQ GEMTERAYOSLLPODYSTONNACFYGENDLEKSVFRIRTDLKNKKHMLERMWSFFESHV LMSGSGATLFVCYLEELEQDSKVSSQIHSLIKQTQGIPVSRLYREPHWYSLKQSTYKNSP

CPn_0956 1089545 1090909

CREDOS 1089545 1090909

CT885 hypothetical protein

LWHEMILPPYSYSLKIGAAVLFFCSILHTFLTPWLYTLCQSYEHKKLVFPECWKRYARL
SELERILSRVEIVFFLWAVPLFFWFLYTEGYRISMAYFNSRNYGFAVFIMVILLLESRP
IWFAELVLSSIAKLGKTSPKSWMYTLMIAPPLSCLLKETGAMIIGATLLMRFFYYFSP
SRFRAYATMGLLFSNISIGGLTSYVSSRALFLIFPALKWEHSFFLSHFAWKAIVAILIST
TIYVFIFRKEFKKFPDIPSDKOPSVEKVPWWIICVNIIFVGSILBRSTFLFWGALLLFY
LGERKFTLTYODPINLSKVCYYGLFYAGLVVFGDLOEWWYLNILMGGLSPGYMTVSYTLS
IFLDNALVNYLVHNLSVATDCYHYLVVAGCMAAGGLTLVSNIPNIVGYLILRSAFPSSTI
HWEWLFLGALGPSIISLGVFWLLKNVPEFLYCFFR

TOPHEO957 1093812 1090963

ide/ppr-Insulinase family/Protease III

KIMTRNCKMFWKLLCPILLCTSLSITSCEOFKVVPNQCPLQVSTPAAADQKIEKIICSN
GLEBLIISDPNLPTSGAALLVKTGKNADPEEYFGMAHFTEHCVFLGNEKYEVSGFFGFG
SENNGVKNAFTYPNKTVFVFSVEHSAFSDALDQFVHLFINPKFRQEDLDREKYAVHQEFA
AHBESDGRWHRIQQLVAPQGHCAFAFCGGNASTLTPVTTEKMAEWFKLHYSPENMCAIA
YTSAPLSKAKKQFSKIFSQIPRSKNYERQEPFLPSGDTSSLKNLYINQAIQPTSNLEIYV
HIYESSHPIPLGCYKALAEVLRNESKNSLVSLKNEQLITDLDVEFFRSSLNTGEFYISY
ELTEKGGKHYSQVIDSTFGYLRYIQEHGIPNYTLESISTINTGEFYISY
GLYSIGNEDLSTYPYHSLVYPKYSSEDESALLNLVSDPEOARFVLSSKNSEHWEEATQWHD
BIEDMTYWKAI DGWODGWOSL WE BLALPKDNI PIPKFVTLI GYWHLLKKGEFFPABALS IVSIGNEDLSTYPYHSLVYPKYSSEDESALLNLVSDPEQARFVLSSKNSEHWERATOGHD
PIFDMTYVKALDGVODYGKVOSLKPIALPKPNLIP IPKEVTLDGVHLLKKGEFPFAFALS
YQDDKLTLYHCEDHYYTAPKLSSQIRIRSPQISRSSPQFLVATELYCLAVNDQLLREYYP
ATQAGLSFTSALGGGIDLRYGSYTTIVPALLNSILTSLPNLEISYETFLVYKKQLELY
CGALLNCPVRSGLDELASQVMKETYSNTTKLSALEKLSFSEFQAFASNLFNSVHJEVMVL
GNLSEQOKKDYLEMLQVFTASRSSHATKFPYVELQSQEISEIHHDYPLTANGHLLLQDK
SSPSIGKVCAEMLFEWLHHITFEELRTQQLGYMVGARYREFASRPFGFLYIKSDAYSP
EELLAATSLFLNKVSASPEKFGISQEKFANIRKAYINKILEPEHSLDMMNSALFSLAFER
PFVEFSTPDLKIAIAETLTYEEFLKYCQCFLSNELGTQTSVYIRGTQKTS

1094803 1093793 CPn 0958 CPH_0958 1094803 1093793
plsB-Glycerol-3-P Acyltransferase
tyrairwgesrylryafdnoylpeplyokfsvfhonyidaatkkaaadolevlclowkv
iiedlkniffippyhkkirapiolffilsidffslviddknskilalhrukeieeylargd
nvvllanhofecopolmyyalgktheelmennifvagdrvtsdplargfsmcolllciys
krhiatpelereeklihnoksmoilktilneggkfirvapaggrorkmægglvpsefsp
egievfrllakasnotthfypfalktydilppppkienaigegraiyfapvffnfgaelf
fdalcokeelihcdkhaortlrækvfsivknlyeel

CPO_0959 1096376 1094799
cafe-Azial Filament Protein
ACKYGICTRKVMENEILLNIESKEIRYAHLKNGQLFDLTIERKKVRQLKGNIYRGRVTNI AQLYGITTRKWMENTLLINIESKEIRYAHLKNOQUEDLTTERKYVRQLKGNIYAGRYVNI LRNIYGJAFINIDERENJETHIGODI LENDKKERGMFOMDVDALFERASGAPLIGSGEAPIE EFLKLDJPVI.VOVVKETIGJEKGRAFITSNISIEORYLVLLPNJPHRGVSRKIEDPHMREQL KQLIRJERMENMELIGTRASTTAJSTEAL INBAHOLLUTWKFILEKFYSTEQFELUTSET DILLKKAY TPOT DIMYNKHALI IDDYATYOKKHIMLKKYSPDAJIKIEYYNDSIPMFERSNIE KELIKATHKKIMJSKAYLEFDKTEAMIFTIDVNSGRSTOJGESOVEETLVQINIEAREIA POLIKHMIVALIVI IDFIDMKSRKNORRVLERLKEHMKYDAARCTILSMSEFGLWEMTROR NRESILMJTLIFTLICFYGRANAI IKTPESVVIETERDLKYJINKENSHLCUVHPETASYM KÕEHDDHEMINLAKQUKAKLQINTSDSVHINHYQFFSUTTOESIDL 1097102

CPn_0260 1 10-7102 CT899 hypothetical 2 10-7102 SUDLVSYLSNPORALVISKOFSMCVDNLKLYIFRLKUPODTERIGYSISPEYIREKGE SUDLVSYLSNPORALVISKOFSMCVDNLKLYIFRLKUPODTERIGYSISPEYIREKGE EELLNSPIEVEGSLGRIDSDOWILDISKTOJGLECPVCNNFFSHSVCLPDLQRVISHDE VGSCVFDCRPLIRQELLLESDCFEECSCOGPPERKNILKFLEDRKKHEGNSPFEYL

1097106 10/17277

rl32-L32 Ribosomal Profein KUNDMANAGHA TAG TURANER SECUTION OF SUBJECT OF A SECUCIONAL STREET

CPn_0952 1097301 1099275
plsx-fa/Phospholipid Synthesis Protein
ILSDFMEVQIGIDLMGGDHSPLVVMQVLVDVLKSQSSTIPFAFTLFASEEIRKQIQEEFI
SDLPQEKPKIISAENFYAMEDSPLAAIRKKSSSMALGLDYLQEDKLDAFISTGNTGALV
TLARAKIPLFPAVSRPALLLVCVFTMRGHAVILDVGANISVKPEEMVGFARMGLAYRQCLG
DSKIPTIGLINIGSEERKGTEAHRQTFFMLRETFGEAFLGNIESGAVFDGAADIVVTDGF
TGNIFLKTABGVFEFLORILGDKLEADIQRRLDYTFYPGSVVCGLSKLVIKCHGKACGSS
LFHGILGSINLAQARLCKRIUSNLI

CPI_0963 1698374 1103224

pmp_21-Putative Ouyer Membrane Protein

TPLERKVAMVAKKTYRSVSSFSHSVIVALISAGIAFEAHSLHSSELDLGVFNKOFEEHS

AHVEEAOTSVLKGSDPVPSCKESEKVLYTOVPLTOGSSGESLDLADANFLEHFONLFEE

TTVFGIDOKLVWSDLDTRNFSOPTOEPDTSNAVSEKISSDTKENRRÜLETEDPSKKSGLK

EVSSDLPKSPETAVAMISEDLEISENISARDPLQGLAFFYNTTSSQSISEKDSSFGGIK

EVSSDLPKSPETAVAMISEDLEISENISARDPLQGLAFFYNTTSSQSISEKDSSFGGIK

EVSSDLPKSPETAVAMISEDLEISENISARDPLQGLAFFYNTTSSQSISEKDSSFGGIK

EVSSDLPKSPETAVAMISEDLEISENISARDPLQGLAFFYNTTSSQSISEKDSSFGGIK

EVSSDLPKSPETAVAMISEDLEISENISARDPLQGLAFFYNTTSSQSISEKDSSFGGIK

EVSSDLPKSPETAVAMISEDLEISENISARDPLQGLAFFYNTTSSQSISEKDSSFGGIK

EVSSDLPKSPETAVAMISEDLEISENISARDPLQGLAFFYNTTSSQSISEKDSSFGGIK

EVSSGLNSGGAFCKSFVSNNENTALWENOALSGGAISASDDIDGORCSAIEFSRGNS

LIALGEHIGLTDFVGGGALAAOGTLTLRNNAVVCVKNTSKTHGGAILAGTVDLNSTISE

VAFKONTAALTUĞALSANDKVI IANNFGELIFEONEVNHGGAIYCGCRSNPKLEQKDSG

ENINIIGNSGA FFEKNRASVLEWTYQAEDVAGGGALMGHNVLLDSNSONIOFIGNIGGS

TFWIGEYVGGGALLSTDRVTISINNSGDVYFKGNKGQCLAQKVVAPQETAPVESDASSTNK

DEKSLNACSHĞDHYPPKTVEEEVPPSLLEEHPVVSSTDIRGGGAILAGHIFITDNTGNIK

PESGNLGGGEBSTVCDLAIVGGGALLSTNEVNVCSNQNVVFSDNTTSNCCDSGGAILAKK

VDISANHSVEFVSNGSGKFGGAVCALNESVNITONSAVSFSKNRTRLGGAGVAAPQGSV

TICGNQGN/AFKENFVFGSENGRSGGAILANSVNITONAGDILFVSNSTSYGGAIFV

GSLVASEGSNRRTLTITNSGDILFAKNSTQTAASLSEKDSFGGAIYTONLKIVKNAGN

VSFYGNRAPSGAGVQIADGGTVCLEAFGGDILFECNINFDSSFNATHLGGNSXIVELSA

VQDKNII FODALTYERITIRGLPKOVSYPLSAPSLIFNSKPQDDSAQHHEGTIRFSRGVS

KIPQIAAIQEGTLALSQNAELWLAGLKQETGSSIVLSAGSILRIFDSOVDSSAPLFTENK

EETLYSAGVQINMSSPTENKOKAVDTPVLADIISITVDLSSFVPEQDGTLPLPPEILIFN

EETLYSAGVQINMSSPTENKOKAVDTPVLADIISITVDLSSFVPEQDGTLPLPPEILIFN

EETLYSAGVQINMSSPTENKOKAVDTFVLADIISITVDLSSFVPEQDGTLPLPPEILIFN

ETLYSAGVQINMSSPTENKOKAVDTFVLADIISITVDLSSFVPEQDGTLPLPPEILIFN

EETLYSAGVQINMSSPTENKOKAVDTFVLADIISITVDLSSFVPEQDGTLPLPPEILIFN

EETLYSAGVQINMSSPTENKOKAVDTVWGSGGGVCACORGFGFKHHLTGYAGLGDTOLVE

DFLIGGCFSQFFGKTESQSYKAMIDVKSYMGAAYAGILAGPWLIKGAFVYGNINNDLTTD

VYDTLGISTGSNIGKGFIAGTSIDVRYINPRFFISAIVSTVVPFVEAEVVRILDAL

AYSKKSYGVDIPCKAWARLSNNTEWNSYLSTYLAFNYEWREDLIAYDFNGGIRIIF

CPN_0964 1104812 1103301
No robust homolog present in Genebank/EMBL as of 11/7/98
QSILESIIKYFYLIHNSKMHMSNPISLFSPAELIAKYNLIPKTSPIYPRRTELIILEENA
COTRITNVAQVLHPSSLFSMSKKILNPCGCSGGPLCWVILNILAFIITSVLFIILLPVNL
VAGGRIFMPLPPKKIVEDLSEPTTEETNEVIQPFIFALQALLFEDNKLRSFKIVEGSVG
KAPLPNPFLNRLVAISPQESQEAMRKIPDLCSQLKKVLKSLGVLTPEWKHMLKYFEGLKN
HHDSNPDKKTFPILIKLLIEALTGKSSLPKTFSTKEKNQAALFIASSCKTCKPTWGEVIT
KSLNRLYSIANEGDNQLLIWVOEFKERELMSIQDGDAEEYRFAAQQHGERYTEAIQVL
RNESAAKLQWHVINTMKFFHGKNLGLVTEHLQDTLGALTLRQTTVDTHQGREDADLSAAL
FLNRYLNSGNQLVNSVFKSMQKADPETKALIREFALDILYASLRLPQTSAHTEVFSTLLM
DEPTYFEDNKACIAVILVJKKIFI. DPETYEPNKACIAYLLYVLKIIEL

CPn_0965 1106769 1104925

lpxB-Lipid A Disaccharide Synthase
KGFSFSKVGLMMIPSGLVYLLYPLGFLASLFFGSAFSIQWMLSKKRKEVYAPRSFWILSS
IGATLMIVHGTIQSGFVYVLHVINLITYLRNLNITSSRPISFRATLVLMALSVVFVTLP
FLYVNMEMMASPNIFHLPLPPAQLSWHLIGCLGLAIFSGRFLIQWFYIESNNTKDFPLLF
WKIGLLGGLLALVYFIRIGDPINILCYGGLFPSIANLRLFYKEQRSTPYLDTHCFLSAG
BASGDILGGKLIQSIKSLYPNIFRWGVGSPAMROBGLQPILMMEEPQVSGFAEVLGSLFR
LYRNYRKILKTILKHKPATLIFIDFPDFHLLLIKKLRKHGYRGKIIHYVCPSIWAWRPKR
KRILEQHLDMLLLILPFEGGLFKNTSLETVYLGHELVEISDYKEQASWKEKFLNSDRPI
VAAFFGSRGGISRNLRIQVOAFINSSLSGTHQFVVSSSSAYYDEIIEDTLKAEGCOHSQ
IIPMNFRYELMRSCDCALAKCGTIVLETALNQTPTIVMCRLRFFDTFLAKYIFKILLPAY
SLPNIIMNSVITPEFIGGKKDFHPEEIATALDLLNQHGSKEKQKEDCRKLCKVMTTGQIA
SEEFLKRIFDTDLPAV

1108055 1106748

SEEFLKRIFDTLPAV

CPG_0966 1108055 1106748

pcnB_2-Polya Polymerase
LLITIIMVCENNILSGRGLELLKKKSNITLTFTIYSVSNHNIKLKDFSPHALSVIKTLRK
AGY1AYIVGGCIROLLLNTTPKDFDISTSAKPEEIKAIFKNCILVGKRFRLAHIRFSKQI
IEVSTFRSGSTDEDVLITKDDHLWGTPEEDVLRRDFTINGLFYDPEHEEIIDYTGGVMDLR
NRYLRTIGDPFTRFKQDFVRMLRLLKILSRSPFTVETQTQEALIACRQELIKSSQARVFE
ELIKMLNSGAAKNFFQLLIENHLLEILFPYMDKAFRLMPALEEQTATYLKALDDKILKKE
AEYDRHQLMAIFLFPLVMFNVRYKHQKHPYLSLTSVFFVIKNFLEQFFADSFTSCSKKNF
ILTALILOMQYRLTPLIPTKKALFFNKKLLHHTRFLEALSLLEIRSIVYPKLDKVYVAWI RHHOTLKCKKDSHSQK

1108431 1109885 CPn 0967

CPT_0967 T104431 T109895

PTAYKPAFICACRSEKIRRIGIDFRRNMOGSVEKLFSTEDVRGRANFEPMTVETTVLLGK
AVARVLRESREXERRIVAVSKDTREXSYMFENALIAGLINSM/IETIALDFIPTPSVAFITR
AYRADASII RÖSZENNEYRRONG IKTESLESSFK ISOVILEGF IETIMVSEADPSFELDEDHAVGK
NKRVIDAMGRYVEFVKATFPKGPTIKGLK IVLESCAIGASYKVAPSVFEELDAEVICYGGE PTGTITINERGALFIOVIÇKAY FERQAHLGIALD JODDE LIMVDEKGHIVA DMILGIAC GDLKKRGALPHNRVVATINTNEGVIKYLEGIA ILI JODDE LIMVDEKGHIVA DMILGIAC SCHMIFELDYNTNID UVGALQVIRIM FEDERMINDDITAFI VKORYPLI NVAVREK I PLET LBFTEB.LFBDAGDVFTB.SERTFFBAGGLEDTGRAWAEGHKRHÖADGFVKVFVDATDVEFV

1104889

CPn_0968 1109889 1111721
qlm:-Glucosamine-Fruetone o P Aminotransferance

DRMCG: FFCYLCHOODGVS: IVLEGLAKLEYRGYDSAGLAA QERE: FTASV: IGHTRWATHGVPTE: INAHPHVDEGRSCA SFASDTDSE: IVQLFSLYYQESQDLVFSFCQTLAQLRGS FIRKTVGRVOELSNLF I ENFKELRRELTAQGI LIHKDHPHTILCASQES PLILGLGKEETFIASDSRAFFKYTRHSQALASGEFAIVSGGKEPEVYNLELKKIHKDVRQ ITCSEDASDKGGYGYYMLKEIYDQPEVLEGLIQKHMDEEGHILSEFLSDVPIKSFKEITI ITCSEDASDRSGGYYMLKETY QUPEVLEGL LOKINDEEGH LISEFLSOVPI KSFKETTI VACGSSYNAQYLAKY I ESLUSTPVI EVASEFRYRRPY IGKDYLG ILISOSGETADTLA ALKELRRRNI AYLLG ICNVPESAI ALGVDHCLFLEAGVE IGVATTKAPTSQLLLLVFLGL KLANVHGALTHAEDGSFFOGLOSLPDLGGKLLANES LHSWAQPYSYEDKFLFLGRRLMYP FYMEKARKERETAY TEARAYT SCHRIBE FALL TYTTV I TAR JODI TYTEME INMMEWY ARHAIDZIA LAGESCHSULTAVANDAZ FEVEYZHEL AARSTAVETYGVAMAZAMALASSMEIDG PRNLAKSVTVE

CPn_0969 1111803 1112999 tyrP_1-Tyrosine Transport_1 VYVMSNKVLGGSLLIAGSAIGAGVLAVPVLTAKGGFFPATFLYIVSWLFSMASGLCLLEV VYYMSNKVLGSSLLIAGSALGAGVLAVPULTAKGGFPATFYIYSWLFSMASGLELLEV
MTMMKESKNPYNNLSMAESILGHVGRISICLVYLFLFYSLLIAYFCEGGNILGRYFNCON
LGISWIRHLGPLGFAILMGPIIMAGTKVIDYCNRFFMFGLTVAFGIFCALGFLKIQPSFL
VRSSWLTTINAFPVFFLARGFGSIIPTLYYYMDKKVGDVKKAILIGTLIPLVLYVLWEVV
VLGAVSLPILSQAKIGGYTAVEALKQAHRSWAFYIAGELFGFFALVSSFVGVALGYMDFL
ADGLKWNKKSHPFSIFFLTFIIPLAWAVCYPEIVLTCLKYAGGFGAAVIIGVFPTLIVWK GRYGKOHHREKOLVPGGKFALFLMFLLIVINVVSIYHEL

CPn_0970 1113452 1114648 tyrP_2-Tyrosine Transport_2 vyvMSNKVLGGSLLIAGSAIGAGVLAVPVLTAKGGFFPATFLYIVSWLFSMASGLCLLEV NTWINKSKNIPUNILSMAESILGHVÖKISICLVYLFLFYSLLIAYFCEGONILGRVFNON LGISWIRHLGPLGFAILMGPIIMACTKVIDYCNRFFMFGLTVAFGIFCALGFLKIDPSFL VRSSWLTTINAFFVFFLAFGFOSIPTLYYYMDKKVGDVKKAILIGTLIPLVLYVLWEVV VLGAVSLPILSOAKIGGYTAVEALKQAHRSWAFYIAGELFGFFALVSSFVGVALGVMDFL ADGLKWNKKSHPFSIFFLTFIIPLAWAVCYPEIVLTCLKYAGGFGAAVIIGVFPTLIVWK GRYGKQHHREKQLVPGGKFALFLMFLLIVINVVSIYHEL

1114693 1115415

yccA-Transport Permease SCHOOL FERMINATE FERMINATE FOR THE STATE OF

CPD_0972 1116377 1115430
ft%YECell Division Protein Ftsy
RCINSLLFPSYLVSFLLLQLTLLLAMFKFFRNKLQSLFKKNISLDLIEDAESLFYEADF
GTEMFELCARLERTKKADASTIKDLITVLLRESLEGLFSQASQSSQTRPIVSLLLGTNG
SGKFTTAAKLAHYYKERSESVMLVATDTFRAAGMDQARLWANELGCGFVSGQPCGDAAAI
AFDŠTQSAIARGYSRVIIDTSGRLHVHGNLMKELSKIVSVCCKALEGAPHEIFMTVDSTL
GNNAIEQVMYFHDVVPLSGLIFTKVDGSAKGGTLFQIAKRLKIPTKFIGYGESLKDLNEF

1116346 1117527

"SIGG-Succinyl-Coa Synthetase, Beta"
EGKSKELFMHLHEYQAKDLLASYDVPIPPYWVVSSEEEGELLITKSGLDSAVVKVQVHAG
GRGKHGGVIVAKSSAGILQAVAKLLGMHFTSNQTADGFLPVEKVLISPLVAIQREYYVAV IMBRIGHRCPVLMLSKAGGMDIEEVAHSSPEQILTLPLTSYGHIYSYQLRQAYKFMEWEGE VMHQQVQLIKKLAKCFYENDVSLLEINPLVLTLEGELLVLDSKITIDDNALYRHPNLEVL YDPSQENVRDVLAKQIGLSYIALSGNIGCIVNGAGLAMSTLDILKLHGGNAANFLDVGGG ASOKOJOEAVSLVLSDESVKVLFINIFGGIMDCSVVASGLVAVMETRDQVVPTVIRLEC NVE<u>IG</u>KEJVQOSGJPCOFVSSMEEGARRAVELSM

1117523 1118422

CPR_4974

1117523 1118422

"sufcD-Succiny1-CoA Synthetase, Alpha"

VCRERRYMFHSLSKNTPIITOGITGKAGSFHTEQCLAYGTNFVCGVTPGKGGTLWLDLPV

YDSVLEAKCATCCRATMIFVPPYAAEAILEAEEAGIELIVCITEGIFVRDMLEVARVMD

NSTSGLIGPNCPGIIKNGECKIGIMPGYIHLFONIGVVSRSGTLTYPEAVWGLTDQKICOS

ICVGLIGGDPLNGTSFIDVLQALEEDPYTELILMIGEIGGSAEEEAAAWIQAHCYKPVVAF IACATAPKGKRMGHAGAIISGNSGDAKSKIQVLRESGVTVVESPAHIGKTVDAVLRAKEL

CPn=9975 1119038 1119637
No robust homolog present in Genebank/EMBL as of 11/7/98
GIEEQVALSIAIKILKIILALILFPLVLLAWVIRYQLHANFHCSVVPFPGFSVNQAYKCS
EAKIEEMLDLLDLETLEWSSRCLRQDMTFANRLEEELIGELRVSETEEL/ISLGGKRNLVR
LLLTHFFNPPKRSRVESVGHEVVFPVFDRLKREEELIGDGPITRSNEELWALLDHGTARG IHKTLWFSIFFKYLTQIELF

CPn_0976 1120079 1121185

No robust homolog present in Genebank/EMBL as of 11/7/98

ILMLVYCFDPSVPTSPEHRIMAALDRWFFLGGHRARILITLEGNHKRARGENNSISTVEKI

KKLISYLLIPIVLIALLIRCFLHSRFKCMWKCDSISDARVPHDVOPFNDFGLDFLOFNDGERLN

IWKNRRVYSGIDVLMVPVDYLRSOFFGFKEIPEAIRCENVVSGOFSEESKTSYLRAMLT

DIVCYILSLDETYWTNVILKIRAMCITFESFFGKEADPNYSGRVTHHYFDESWKALARHV

LGEGNMVNRLDEALIRTEKRGKEGECITKOFLKDYCKKHLEVMSCPDFIESLVDEKIREF

RCPSILNSAVCDVIDRKCOEHLLKAIINEANRRLFGMKNSFTMRGNQVLFYTIFSPPKL

PRAASSQVF

CPn_0977 1121329 1122402
No robust homolog present in Genebank/EMBL as of 11/7/98
LYINGFANILKSSFLMEVYSFSEVETSFOHRVMAALDMWFFLGGRRLKVVSLDSCNSGQ
ACEEYVPISTTEKVLKILSYLLIPIVIIALLIRYLHSNFTAKVSOKFMLKHILDLGIDIK
SFILPGSHVNTMDSATLFKAIRLECKRVDVEYHRUHSSDKVVFYIPAQKLPDDLRLTHWL
PEKETRKTEYVRHMLAHVMGYLTSQGKRELDOVYCDSRSTSTSLGAEKVLOVRFIDHPOSO
GEFURLLNENITTKGSEDKEVVQSDLFDMAFOGMMPQFISVIGSPTFSEELVHEMSQKLD
LDCIYPEDDEFEQKFLNTLLKAVLHHGFEGISVASMGVIFLICPDSLALQIPFLRNQK

112265+ 1127603

CPD_0978 112265 (112760)

No cobust, homolog present in Genebank/EMBL ad at 1177798
KYFEMEVYGEHEAVRIGEGHRVMAALDAMEFLASHRLEVAVILDIGENGGMAYGELVSTSTT
EKVIKLEDYILLVETVETALLERCLLHON REDVEKERWIKETRELGIDTESCKLEGSYVNO
VISHETWEEKDRICKERRETOVYTEHERSKWAVFETVEGKETERREDYMEGKETRKROYV
RIMMIDIIV TOTTITIST GEWILDYTEHERSKWAVFETVEGKETCHONGELGGEVORLLHEE
GATROGGOKEVILGHIVSDITCOCWWKFLEVTQSFAFTFELVEEVGGKINLDFLCLEKAN
TELOGERRETELWAVVIHRGIBTOVITKEVGAGLITYTEATOLGTPEGRS

1125/143 CPn_0979 htrA-DO Serine Protease

htra-DO Serine Protease
GIDMITMOLRSWLAVLVGSSLLALPLSGG, VVKKESRVSELPODVLLKEISGGFSKVATK
ATPAVVYIESFPKSQAVTHPSPGRRGPYENPFDYFNDEFFNRFFGLPSOREKPOSKEAVR
GTGFLVSPDGYIVTNNHVVEDTGKIHVTLHODGKYPATVIGLDPKTDLAVIKIKSONLPY
LSFGNSDHLKVGDWAITAIGNPFGLQAT/TVGVISAKGRNQLHIADFEDFIGTDAAINGN
SGGPLLNIDGDVIGVNTAIVSGSGGY/GIGFAIPSLMANRIIDQUIRDGOVTRGFLGVTL
WILLIAMIYKLEFFG WINTEWY/ WARDE WERE, ELLI ANTERGADE, MERINAVI
LMESETRIVLKYSREWYIESFTT/ ARKED SLAALS WISS JURGJESTAPPLLIAIE tkgiliisvepgsvaassgiapgqlilavnrokvssiedlnrtlkdsnnenillmvsogd VIRFIALKPEE

1126/988 1125504

similarity to Saccharomyces serevisiae hypothetical 52.9KD

*similarity to Saccharomyces serevisiae hypothetical 52.9KD protein
FYMLINHAKKHAKPYVLIFFSTKDKLSYCDIIFNNCSGKPMNLDSKHFDINSANFLEEFAK
FISFPSISADSDHLQDCENCAHFLVDHVNKIFDVELWETPGHPPIIYASYKSEDPLSPTL
MLYNHYDVQPAQLSDGWXGDPFILREENGNLYARGASDNKGCCYTITALQHYYESGONF
FUNIIWLIEGEEESGSIALFTWLEKKKEALRADYLLIVDGGFLSEKHPYVSIGARGIVSM
KISLEEGNKDMHSGVLGGIAYNTNRALSEILSSLHHPDNSIAIEGFYDDLALPSDSDRPD
LPKSDTLRECEENLG/RPGGYEASYSPEESALRPTVEINGISGGYTGGFKTVIPYRATA
YLSCRLVPNQDPDKWAHQVIHHLKQOVPSSLKFSVEILDGGSRGWRSSANLPIVKVLGEI
YSDLYNEECLRLVMPATIPIGPLLGEAAQTSPIICGTSYLSDDIHAAEEHFSMDQLKKGF
LSICQLLDKLPKKE

1127019 1129952

Zinc Metalloprotease (insulinase family)
VTESMKAGDTYRNFIIKSCKDLPEIESKLLEAEHKPTGASIMMIVNNDEENVFNICFRTC VTESKKAGDT/RNFIIKSCKDLPEIESKLLEAEHKETGASIMHIVNNDEENVENICERTC
POTSNGVAH/LEHMVLCGSENYPVRDPFFSMTRRSLNTFINAFTGPDFTCYPAASQIPED
FYNLLSVY/DAVFHELITKQSFLQEAMRYEFNSENHLCYTGVVFNEMKGAMMSGEARLSE
ALNAAIF/SVYTGVNSGGEPREIVTLSHEDVRAFHQSGYSINRCLFYFYGNIKPSRHLDF
LEEKLLBOATKLEKQAVSVPLQKRFKEPVRNILTTYPUDHQEEDKVLFGISWLTGSILEQO
ELLALH/LEIILMGTDASPLKSRLLKSGFCKQTEMSIENDIREIPMTLVCKGCSPAGAQK
LEALIFASLEEIIREGISENIVEGAVHQLELSRKEITGYSLEYGLSIFFRSGLLKOHGGS
AEDGERIHSLFSELRNSLKNSDYLAKLIRKYFLDHPHFARVILLPDTELVAKDNKDEQQL
LLSYSEKLTDENKEKIQQNVRELTESQEQKEDLNGILENLALDKVPTSCKEFFLIKEGLS
OGBVLHHECFTNDIVFIDVVLDIPPLSGEELPWLRLLVFLMLQLGCGGRSYKEHLEFLLE
HTGGVDVSYDFSPHANNSFLSPSVSIRGKALSSKSEKLCGIVSDMLTSVDFTDIFRIRE
LLMGHNEALTNSVNSPMSYAVSMACSGNSITGAMSYLTTGLPYVKKIRELTKNFDQNID
FAVVILORLYTKCFSGRQIVISGSANNYQQLKDNKFFGLLDYLVLIVIEPPENPSINLY
TSRGLHIPARAFNALAFPIGDIAYDHPDAALTVASEILDNVVLHTKIREQGAYGSGA
AANLSRGSFYCYSYRDPEIATTYKTFLKGVSEIASGNFYKEDIYEGGALGVVQGLDMPVAP GSRASVAFYRLKSGRIPVLRQAFRRSVLEVTKEHICMVMDKYLESTVQETTLISFAGEEM

1131215 1129962 CPn_0982

LRNNVLTLDKDFPIVPAI

CPH_0982
yigh family
KKELASVMNLPVSLACLLLSGCVFFLGVFVSSSLYARKKRAFLEKIQKLEHENQLLQTSL
NLSRHQEQLIEDFSNRLALSSKKLIKDMKEEAQNYFGDTSKSFQSILSPIQTTLTTFKQS
LETFETKHAEDRGRLEGUISOLLAVEKKLEHETHVLTDILKHPSGRGMGEJQLERILEL
AGMLKYCDYDSQTTSAQGAFRADIIIRLPQDRCLIIDAKAPISDSYFSVEEIDKGDLVDK
IKEHIKTIKKSKYWEKFHQSPEYVILFILPGESLFNDAIRLAPELMEIGASSNVILSSPLT
LLALLKTIAYMMKQENLQKQIQEVSLLCKKELHRRLQVVFTHFQKIGKNLNQTVOSYNDMT
SSFQYRVLPTLRKFEGLETSSSHQIEEPTPIESLATSFPHTCDIDTNLAVIESLEKQD

1132045 1131206

CFI_0783

PSSA-Glycerol-Serine Phosphatidyltransferase
KNPLCYEQKKLWGIDMAGLDLEARGKRRVYTDNAITAFGLCCGLFIIFKSVLRTSSSVEL
FHRLQCLSLLLISAMIADFSDGAIARIMKAESAFGAQFDSLSDAVTFGIAPPLIAIKSLD
GIVYGNFFSSLLLITSILTISLGGVLRLVRYNLFSQKTVDVSKPYCFIGLPIPAAASIVS
LALFLASDFFPDLPAQLRVGLLSFALLFIGGLMISPWKFPGVKHFRFNVSSFLLVVTIGL AACLFFSGLVDHFVEVFFLVSWLYTLVGFPIFSIIYRKKS

1132370 1135510

CPn_0984 1132370 1135510

"nrdA-Ribonucleoside Reductase, Large Chain"
GKVMVEVEEKHYTIVKRNGMFVPFNODRIFQALEAAFRDTRSLETSSPLPKDLEESIAQI
THKVVKEVLAKISEGGVVTVERIQOLVESQLVISGLQDVARDYIVVRDORKAERGNSSSI
IAIIRRDGGSAKFNPMKISAALEKAFRATLQINGMTPPATLSEINDLTLRIVEDVLSLHG
EEAINLEEIQDIVEKQLMVAGYYDVAKNYILYREARARARANKDQDGGEFVPQEETYVV
OKEDCTTYLLRKTDLEKRFSWACKRFPKTTDSQLLADMAFMNLYSGIKEDEVTTACIMAA
RANIEREPDYAFIAAELLTSSLYEETLGCSSQDPNLSEIHKKHFKEYILNGEEYRLNPQL
KDYDLDALSEVLDLSRDQOFSYMGVONLYDRYFNHLHEGRRLETAQIFMMYSMGLALNEG
EQKNFWAITFYNLLSTFRYTPATPTLFNSGMRHSQLSSCYLSTVKDDLSHIYKVISDNAL
LSKWAGGIGNDMTDVRATGAVIKGTNGKSGGVIPFIKVANDTAIAVNCGGKRKGAMCVVL
ENMHLDYSDFLELRKNTGDERRFTHDINTASNIPDLFFKPLEKKGMMTLFSPDDVFGHLE
AYGLEFEKLYEEYERKVESGEIRLYKKVEAEVLWRKMLSMLYETGHPWITFKDPSNIRSN
QDHYGVVRCSNLCTEILLNCSESETAVCNLGSINLVEHIPNOKLDEEKLKETISIAIRIL
DNVIDLNFYPTPEAKQANLTHRAVCLGVMGFODVLYELNISYASQEAVEFSDECSEIIAY
YAILASSLLAKERGTYASYSGSKWDRGYLPLOTIELLKETRGEHNVLVDTSSKKDMTPVR
DTIGKYGMRNSGVMAIAPPATISNIIGVTOSIEPMYKHLFVKSNLSGEFTIPNTYLIKK
KELGLWDAEMLDDLKYFDGSLLEIERIPNHLKKLFLTAFEIEPEWIIECTSSRQKNIDMG
VSLALYLAEPDCKKLSNNYLTAWKKGLKTTTYYLROQAATSVEKSFIDINKRGIQPRWMKN
KSASTSIVVERKTTPVCSMEEGCESCQ

1135432 1136571

CPI_0985
I135432 1136571
intdl-Pibonucleoside Reductace, Small Chain*
ISVHKYCGRKKNNPRLFNGRRLRILSITEKRGAKMEADILD/KLKRVEVSKKGLVNCNQV
DVNQLVFIKYKWAWEHYLMOCANNMLPTEVIMARDIELW/KDEDJEDERRYTELINLDFFS
TABSILVONNITVLA IFRHITNIPEAROYULMQAFEEAVHTHTFELVICEGGALDEGEVFNAYN
RRASIEAKDDFQMMLUVDVIJDHIFTVQCOEGLOGFIKNLWGYYLIMAXIFFYGGFVMILS
FHRONYMTGIGEQYQYTLRDETTHLNFJIDLINGIKEENFFWMTTHIQEETVALIEKAVE
LETEYAKDCLPRGITGJESSMFIDYVRHIADRRLERIGLYPIYHGRNPFIWMGFTMDLNK
EKNFFETRYTEYQFASNIGW

14.667121137395

PÄLITYAMMELLÄÄ I KMULLA GERMERITRICKS VET LÄLEL I KRYGT.
PÄLIVETAAN PROPERTIERE PROPERTIERE I KRYGT.
PÄLIVETAAN PROPERTIERE PR

1137483 1139115 VtgB-like predicted rRNA methylase
LENGIFAIGFFMFAYRTLLTHRVVQVSHEIFKTTVVFGDTVIDATCGNGNDSLFLARLLQ GEGRLVVYDIOKEALSNALLLFETHLSEQERSVIEMKEQSHEHILEKDVKLIHYNLGYLP KGNKEITTLARTTEISLEYALNIVRPDDLITVVCYPGHPEGEKETHSVESLAQRLHPKEW CVSSFYVANRCRAPRLFIFQRQGSESSVDKG

anger Old II Anna y Innel y Lawrey Indiana (Architecture Reshintar)
KFFINLINLDQGILKMKEAAPMHFPFPVRRSVWLARYSFFRIGGPANYKKAIHTIEEARE VIRFLHSINYPFLIIGKGSNCLFDDRGFDGFVLYNAIYGKQFLEDARIKAYSGLSFAALG KATAYNGYSGLEFAAGIPGSVGGAIFMNAGTNESDISSVVRNVETINSEGELCSYSVEEL ELSYRSSRFHRQOEFILSATFQLSKKQVSADHSKSILQHRLMTQPYTQPSAGCIFRNPEG TSAGKLIDAAGLKGLAIGGAQISPLHANFIINTGKATSDEVKQLIAIIQSTLKTQGIDLE HEIRIIPYQPKIHSPVSEK

1139552 1139016 CPn 0989

CFBJ2 hypothetical protein LRTSLAVKCVLLTIFWILLVMATLSPEKFSGSPISISKEFPQQKMREIILQMLYALDMAPS AEDSLVPLLMSQTAVSQKHVLVALNGVTKSILEKSQELDLIIGNALKNKSFDSLDLVEKNV LRLTLFEHFYSPPINKAILIAEAIRLVKKFSYSEACPFIQAILNDIFTDSSLNENSLSI

CPn_0990 1139880 1140440 infc-Initiation Factor 3 SVALNRKINRQIRAPKVRLIGSAGEQLGILAIKDALDLAREAGLDLVEVASNSEPPVCKI MDVGKVRYGLIKKEKDSKKAQHQVRIKEVKLKPNIDENDFSTKLKQARTFVEKGKKVKIT CMFRGRELAYPEHGFKVVQKMSQGLEDIGFVEAEPKLAGRSLICWAPGTVKTKKKQEKS

CPn_0991 1140394 1140612 rl35-L35 Ribosomal Protein

KORKNRKSLMPKMKTNKSVSARFKLTASGOLKRTRPGKRHKLSKKSSQEKRNLSKOPLVD KGQVGMYKRMMLV

CPn_0992 1140622 1140996 r120-L20 Ribosomal Protein GKLVMVRATGSVASRRRKRILKQAKGFWGDRKGHIRQSRSSVMRAMAFNYMHRKDRKGD FRSLWIARLNVASRIHSLSYSRLINGLKCANISLNRKMLSEIAIHNPEGFAEIANQAKKA

CPn=0993 1140975 1142030 phaS-Phenylalanyl tRNA Synthetase, Alpha KSPSSHSLGIRISMEMKEEIEAVKQQFHSELDQVNSSQALADLKVRYLGKKGIFRSFSEK LKÖSTDKAKLGSLINDFKTYVEDILQEKSLVILASEQAEAFSKEKIDSSLPGDSQPSGR HILKSILDDVVDIFVHLGFCVREAPNIESEANNFTLLNFTEDHPARQMHDTFYLNATTVL RTHTSNVQARELKKQQPPIKVVAPGLCFRNEDISARSHVLFHQVEAFYVDHNVTFSDLTA ILBAFYHSFFORKTELRFRHSYFFFVERGIEVDVSCECCGKGCALCKHTGWLEVAGAGMI HPOVLRNGNVDPEIYSGYAVGMGIERLAMLKYGVSDIRLFSENDLRFLQOFS

HPOTERION OPPETYS GYAVGMGIERLAMLKYGVSDIRLFSENDLRFLQGFS

TOTO 1994 1142371 1144440

CTB37 hypothetical protein
LFWERRGRIKRSRRNFEQALENLEKLKEISLATSNDSYLINDPARFNORKOTGSSVMENK
PALKAVENYLLEISCVSKSHADKALKESDFLIAGVONVFSFLENQEDLYKSLLDEYSFVT
KAYPBEVKKNIKEVPTYDLSTDEETEEHKEPECFLINILVEVKRDRSYELFYMLDEQDKFY
NDALWOITYKONKLHETVNEGDPLIKTLLWINSEEVENIASSLUVINDMPLELFFYOKLSH
LDIEAVVKVHNAVMALFFSRYEATMVFKSPKKHNIWYFNDFLLFLREAWKDLINDWIDSQ
ERKOTKLLASALSIGIFESKLVFEEASRYLYFNIOTKLENANGKKPLSPGOYLTDAVEEL
HRLESKYPNOGPLFKANDRVLEHESRYDPMILGIDESLGTLKLHGKSIDIISFSPTVG
SSILYANCNEEFLGFLNAKAHRSEVTLVLNIQNRISRKERARSRVIEEALECEHAPYVH
AFSEPPEELLGOLLESIHGDIETFADFFSILOEFHKPLLASSFFLTKELKEFVGSSILKE
KLTALKOIFFAKKKILFROKLLLLHLLSYLIVFKLEITNINDSITVVSKOGLDYVSVFI
AGRÄFFSREAFWDEHSLKLLLTNVLSPTLVARDRLVFVSHIELLSKFVACLKKNRQGFSS
LKSFFKNDIEGWEFTGYLHELTEVSHKHNL

Ph_0995 1145515 1144415
CT0386 hypothetical protein
RMLEWKRHLTREWFALTSLLVLALIFYASIHHSLHTLKGASTAASGASVKLSILYYLAG
ISÉKÁEFLMPQLVAVATTSTLFAMQNKREIILLQASGLSLKSLMHPLLLSGAVIMMVLYA
NFOWLHPICEKISITKENNDRGTTOKEGGKIPALYLKDOTVLYSSIEPKTLTLNNYFWI
KDPKTIYTMEKLAFTTLSLPIGLNYTOFFANDSENLELKEFTOMKEFPEIFFNFYEPPFS
KLFSAGNKNRLSEFFKAIPWNATGLGLSTOVPQRILSLLAGFYYVLISPLACMAAIILSA
YLCLRFSRTPTVTLAYLIPLGTVNIFFVFLKAGIVLASSSVLPTLPVMAFPLIVLFLLTN
YAYAKLQ

CPn_0996 1146592 1145519
CT839 hypothetical protein
AMPILMKVLIFRYLKTAAPCTLSLICISIISSLQE/VAYIAKDVPYDTVLRLMAYQIPYL
LPFILPGSCFVSAFSLFRLSDNNHMTFLRASGAGOSIIMFPVLMVSGAICCLMFYTCSE
LASICRYOTCKEIANNAMTSPALLLOTLOKKENNRIFIAVDHCAKSKFDNVIVALKGNNE
ISHVGIIKŠIIPDTTKDTVKAKDVVFISKLPD9LTESSSPSSQRFYIETLDELLIPKITS
TLFAGKSYLKTRTDYLPMKQLVKQSLKHSHLBETLRRVAIGFLCITLTYAGNIGIHKPR
FRKSIALYFIFPILDLILLIVGKNTKNLPLAFMLFVFPQLVSWVVFAARAYRESRGYA

CPn_0997 1146699 1147664
mesJ-PP-loop superfamily ATP-se
AYKWLSSDLENDOKOLDEFSALDVKKYYLLALSGGSDSLFLFYLLKERGVSFTAVHID
HGWRSTGAQEAKELEELCAREGVPFVLTTLTAEECGDKDLENQARKKRYAFLYESYRQLD
AGGIFLAHHANDOAETVLKRLLESAHLTNLKAMAERSYVEDVLLLRPLLHIPKSSLKEAL
DARGIGYLOPP://DEDERYLRAMMRKKLFPWLEEVFGYHITTFPLLTLGEESAELJEYLEKQ
AQPFFGAATHOD://GGELPCPDCLID/AFLCKWYMKKFFNNAGIAVJRHFLQMVYDHLGRS
GCATLPMRNKIVIIKTJVVVID

1147/11 1150584

GET_0.008

1147/31 1150584

LUBH-ATP-dependent zinc diotaise

LUBH-ATP-dependent zinc diotaise

LUBH-ATP-dependent zinc diotaise

LANIRE IVERDINK IALNOMANIFCORREDVOTQEGGLEVINLEE IDOSHREDEDLQETS

KSUPTIVIKEVING ILWESYLGSBIPEGGYA ESYPSEZIOSIVITEBELVATGRAFPOLINL

BEDDERVITEDRIPEALBYYGEDLYELICKYESPVIGITGSFTLKREENDLYGQVEVSLTQ

ETUPEAAYTLYGOVISTPLNK ISSOLVVSEGGERFSQLIKSVRLYREEWNKYHKLVEARDLN

QAQLEKLRGELSOTVWYFNNQELSSRSLEKQDPEVFSHWFAGAKEEWTAFKFNHGLSFKA

PDOPRNLVLEKTFKSGEPS
PARMLLKONKVTFADVAGI
SELIEIVDFLKNENNFTSLAGRIPKOVLLIGPPOTG
KTLIAKAVSGEADRPFFSIAGSDFVEMFVOVGASRJROMFEOAKRAAPCIIFIDEIDAVG
RHRGAGIGGGHDEREOTLNGLLVEMDGFGTNEGVYLMAATNRPDVLDKALLRPGRFDRRV
VMNLPDIKGRFEILMYHAKRIKLDFTVDLMAVAJSTFOASGADLENLLNEAALLAARKDR
TAVTAVDVAEAADKVLYGKERRSLEMDAEERKTTAVHESGHAVVGLCVOHGDFVDKVTII
PRGLSLGATHFLPEKNKLSYWKKELYDQLAVVLGGRAAEEIFLGDISSGAQDISQATKL
VRSMYCEMOMSPOLGNYTYDERSOCLTDYGGYHEKSYSEETAKTIDTELRMLLDAAYGRA
LEILMBBRABLELMTSMEIREDTLAGERVKEIMDHTAFFEKKRAFLFBERMLFKKSSCEL

CPn_0999 1152859 150766
pnp-Polyribonucleotide Nucleotidyltransferase
QETFMNFQTISINLTEGKILVFETGKIARQANGAVLVRSGETCVFASACAVDLDDKVDFL
PLRVDYQEKFSSTCKTLGGFIKREØRPSEKEILVSRLIDRSLRPSFPYRIMQDVQVLSYV
WSYDCQVLDPDLAICASAALAIBIPPOSNIVAGVRIGGIDNOWVINPTKELASSTLDL
VLACTENAILMIEGHCDFFTEEQVLDAIFFGHKHIVTICKRLOLMOEEVCKSKNLSAVYP
LPAEVLTAVKECAQDKFTELFYKDKKVHAATAHEIEENILEKLQREDDLFSSFNIKAA
CKTLKSDTMRALIRDREIRAGGSLTTVRPITIETSYLPRTHGSCLFTRGETGTLAVCTL
GSEAMAGRYEDLNGGLSKFYLQYFFPPFSVGEVGRIGSFGRREIGHGKLAEKALSHALP
DSATFPYTIRIESNITESNGSSSMASVCGGCLALMDAGVPISSPIAGIAMGLILDDGGAI
ILSDISGLEDHLGMDFKYAGSGKGITAFQMDIKVEGITPAIMKKALSQAKQCCNDILNI
NNEALSAPKADLSQYAPRIETMQIKPTKIASVIGRGGKGIRGIIEETGVQIDVNDLGVVS
ISASSASAINKAKEILGLUGVEVEVGKTYRGRVTSVVAFGAFVEVLPGKEGLCHISECSR
QRIENISDVVKEGDIPOVKLLSINEKGQLKLSHKATLE QRIENISDVVKEGDIADVKLLSINEKGQLKLSHKATLE

1153193 1152891

CFI_1000
TS15-S15 Ribosomal Protein
SAFAAIILRRHMSLDKGTKEEITKKFQLHEKDTGSADVQIAILTEHIAELKEHLKRSPK
DQNSRLALLKLVGQRRKLLEYLNSTDTERYKNLITRLNLRK

1153369 1153869

yfhc-cytosine deaminase
yylelceeklinmekDiffmooafkearkayDoDevPvgcvivkDDkiiarahnsvekLk
DATAMAEILCIGSAAQDLDNWRLLDTVLYCTLEPCLMCAGAIQLARIPRIVWAAPDVRLG
AGGSWVNIFTEEHPFHTVSCTGGVCSEEAEHLMKKFFVEKRREKSEK

h 1002 1153844 1154089

TT845 hypothetical protein
KSAERKVKNKIVTLLDQLYEDQESRLQKLGEEIVPNLTPEDLLQPMDFPQLEGNPAFRFE
EGVLSGIGEVRAAILAALSQEN

1154862 1154092

CT846 hypothetical protein
T\$NKTIHPLLWGPDRQIAGKASMRVIFPDKHNNFPNLSKLLKKLPSVILVTSCIAPFFSY
HINKFFGIPGLLEILALSVKGIQKHHFWOFLTYPLITADSLSLNKDQSFEITQRLLLRNV
LDFFLFYKAIQHLIRKIGAFSVLVVISQQALIIGAVLWGFWALIHSSQSFFGPESIIGGV LTVQIFLDPEKRFTIGPTPLSVSIKWGFLFVLGFYCCILIFSGAFLLLLASMLAIVLAIL FCKKEKIPNPYTTSLRF

CPn_1004 1155418 1154879
CT847 hypothetical protein
HLSIELMSIQPVSNTTKADKVIPDSTKVISDSITINKQSAFYFCISVMLRLSESTTEY
GKSILAVLEDNTIVQQQRVKELINLPLLKVPDLQKKDGSDDEYKNQNEIQAYQSSNQQIS
ANRQMIQQELSSAQQRAQANQKSVNSTTIESMQILQATSSMLSTLKELTIKANLTNSPSD

1155957 1155415

CFR_10US
CT848 hypothetical protein
NRKPVRLMMIIIDPLSAKKPLQAAINVPGTPITGPNTATADDIIAKFSKDSNPLIVTVY
VYVOSULVAQDNLSIIAQELQANSSACTYLLNNQSALYQVVSIPKNKLNDNSSSYLQNIQS
DNQAIGASRQAIQNQISSLGNAAQVISSNLNTNNNIIQOSLQVGQALIQTFSQIVSLIAN

CPn_1006 1156493 1155990 CT849 hypothetical protein TKVNFFIMSITTLGTLPTVNTINSSRPPLEPLNTPKIGAVLFSIYELLLQAIEIRQQTVL TOSOQLNDNTNIQOQLNOETNQIKYAIVSAGAKEDEITRVQNONONYSAQRSNIQDELVT TRQNGQIILSHASTNINIIQOQSSQDSSFIKTTNSIGSTVNQLNKPLG

1156689 1156907

1156089 1156907 CT849.1 hypothetical protein LWYKSLAGEEKDVSGNECNDYPEVFKDDVSAYVLVTCGQMSSEGKIQVEMTYEGDPAVIS YLLTKARDSLDES

1156904 1158223

CFH_1008
CTB4850 hypothetical protein
VLNYSFIGMLKPMYVLSKRLYRWVNQLIKLGDLVKNSRSFSVEWVFISALLLIFGCLGCA
SVVKVSLVPFLLLFSFLAFFLILCFRGKCYALLLGVFYTLYVAKYVVGETLYVSFWLSGL
GVSFLLAFGLFLGCWLAQEEMVKGKEQLRLSEDLAQRSGYEDLLTKSQEKEFLDAR
AQGLDRELTECQELLKAAYQKQEYLTIDLKILADQKNSWLEDYAELHNKYIELVSKNGDV VFPWVAEPSVGESQGSERVDVSRWVSALQEKEESLERLRNEILVEKQRCSDYEHRCQELG LLLQNFTALERRCEELQNLLNQKETQINELHQLVCKSEEKVSVEPSAHAETSCVEEKQYK GLYSQLOEOFLEKSETLSLVRKKLFAVQEKYLTLKKKEELTKQDISFDDISMIQGLLERI EILEEFYSHLEELVSRSLSL

CPn_1009 1159095 1158186 map-Methionine Aminopeptidase YRLLHRYILMKRNDPCWCGSGRKWKQCHYPQPPKMSPEALKQHYASQYNILLKTPEQKAK TYMACOTTARILDELEKASOKOVTTNELDELSGELHKY/DA LAPPHYSSPPFPKTICTS
LIJEVICIPISTENDI PUKUGININ IDVISCTVDGYYGCOSHWM IGEVPETKKKTOAALECL
NDSIGALIKOCTPLESEGA LERARDYYGSVVDGYVGKIYOT EFHENPYVPHYRRSMIP
LAISMI IFTTEPMINAKKESVVDPKNOMEARTCDKGPISAGMENTIA ITETYFE LITLLIND

CHR_1010 LISSEYS 1159067
CTROOF SYPOCHER LOCAL PROTEIN
WHILLIALDILEFVER PROTEING STEVENALLKNEURKKYQRVILREGLEADIALILEFVERGR
OFFOFLD DUVAFOLICGELEFYD I KMMLAFMPEKAKDDYGGTER I FFFLAFFVITGPA
V FYALLEYMEIG I YOREL FFFAM I FAWAFGLEFFLLGGOFFDRLEGNEGLLALERLEGIAL
V EYALLEYMEIG I YOREL FFFAM I FAWAFGLEFFLLGGOFFDRLEGNEGLLALERLEGIAL
V EYALLEYMEIG I YOREL FFFAM I FAWAFGLEFFLLGGOFFDRLEGNEGLLALERLEGIAL DEMOVMEMBERGESTAFNICHYTC

Cfm_1611

1160309 1159903

1154402

CTRS : hypothetical protein EMRLKNYPMIQESFELPOTCILLLASDSLTNILALHH AMFALYGLAL/GLKVLNTPVCATEVVCGIAVTLAGVRA VKORML'/LLRESFFAFI CKEESWI PYKFNMS PSYS PC LOP TALPLMFGPSG

CPn_1012 1162220 1160421
yz:B-ABC crimsporter permease
ALFOLITSYMKKKFIFYFYTVFSLLFLWEMTSRHPPTEGEFCPPPSGIAGGTLOSLPLLL
TOWNFILKA 155 SPELATEL TYGAN THE TWEAT LEGER HAD THE MALALLITUM
HOWHELMAN TO THAT THE TREET THE TWEAT LEGER HAD THE MALALLITUM
HOW HOW HAVE THE THEFT THE TREET THE TREET HAD THE THE TREET TH HIFSGLKIAIGSAGFAAIAGEWVASOSGLGILMLESRRNYEMELAFAGLATLSILTLSLF QITLLIEKLIFSLFRVKRMSLKHKSVAKKALSVLALIPIMLIPWKGNSKSPPDKKNLTSL TITLLIDATENENT FLYKRASLAKTASVAKALSVALITY HILLPANGUNASSPURNITSL TILLIDATENENT FLYKAGVAKGYFKOHGIDLOLOKNITOSSAVPHVLFEQVDMALYHALG IMKTSIKGMPIQIVGRLIOSSLOGFLYRSODPIYKFEDLOKKVLGFCLANSRDLARLLET LNRNGVVPSEVKNVSSDLISPRLLIKKIDFLYGAFYNIEGVKLQTLGMPVKCFLSDTCDLP TGPQLIVFTKKGTKASEPEIVZAFÇKALQESIIFSKDHPEDAFKLYAKETKSIFKNIYQE YLQWEETFPLLAQSQDPLSKDLVDKLLETIIKRYPELASEVAKFSLNDLYNPSLPEEQSV

1162209 1163624

CPH_1013 1162209 1163624
fumC-Fumarate Hydratase
RENSLAHRGNIDMRQEKDSLGIVEVPEDKLYGAQTMRSRNFFSWGPELMPYEVIRALVWI
KKCAAQANQDLGFLDSKHCDMIVAAADEILBEGGFEEHFPLKVWQTGSGTQSNMWNDEVIA
NLAIRHHGGVLGSKOPLHPNDHVNNSQSSNDVPFTAMMIAAVISLKKNLIPALDHMIRVL
DAKVEEFRHDVKIGRTHLMDAVPMTLGQEFSGYSSQLRHCLESIAFSLAHLYELAIGATA
VGTGLNVPEGFVEKIIHYLRKETDEPFIPASNYFSALSCHDALVDAHGSLATLACALTKI
ATDLSFLGSOPRGGLGELFFPENERGSIMPGKVNPTQCEALGWYCAQVLGNNQTVIIGG
SRGNFELNVMKPVIIYNFLQSVDLLSEGMRAFSEFFVKGLKVNKARLQDNINNSLMLVTA
LAPVLGYDKCSKAALKAFHESISLKEACLALGYLSEKEFDRLVVPENMVGNH

1165456 1163732

CPn_1014 1165456 1163732
ychm-Sulfate Transporter
ALASTLGYCIVKVPWAFKNFIPKLYTSIKEGYSFNTFKKDFQAGITVGILAFPFAIAIAI
GVGVSPIQGLLASIIGGLLASAMGGSNVLISGPSSAFISILYCLSAKYGAEALFTVTLLA
GVFLIAFGLTGLGTFIKYMPYPVVTGLTTGLAIIIFSSOIKDFLGLOMGANIPADFLPKW
HAYWDHLWTWDSKFSAVGLFTLLIHIYYFRNYRFYPGVMIAIVTATTLWALLKIDIPTIG
SRYGTLPTAIPLPKIPQLSITKILQLMPDALTIAVLSGLETLLSAVVADGMTGWRHQSNC
QLVAQGVANIGTSLFSGIPVTGSLSRTAASIKSEATTPIAGIVHSIFICFILLLIAPLTV
KIPLTCLAAVLILIAMMNSEIHHFIHLFTAPKKDIVVLLTVFILTITMITITAAVQVGMML
AAFLFMKQMSDLSDVISTAKYFDKDSDFLSKAEVPQNTEIYEINGPFFFGIADRLKNILN
DIEKPPKIFILCMTRVPTIDASAMHALEEFFLECDRQGTLLLLAGVKKTPLADLKRYHLD
ELIGVDHIFSNIKSALLFAQALTNLESKTSTRHLV

CHILIO15 1165550 1166893
CTQT hypothetical protein (possible IM protein)
KNNEWNESFFTSVRVRSKVDHEIILEVTHLKLQLCALFLFGYLAIVFEHIVRVNKSAIAL
ANGGLMMLVCFSHIPMADHMILVEEIADMSQVIFFLFSAMAIVELIDAHKGFSVIVKFCR
IQERTLLUMALIGLSFFLSAALDNLTSIIIIISILKRLVKAREDRLLIGAICVIAVNAGG
AMMEHGDVTTHMWINNKITSWGIIRALFVPSLVCVLVAGFCGOFFLRKGSTLIAKDVE
LOSAPPRSLMIFFIGGSSLLMYPVKACLGLPFPFSALLIGLIGLWALTSDWITHSPHEEDRY
HLERYPHILTRIDISSTFFIGILLAVNALSFANLLTDFSLWMDKIFSRNVVAIVIGLLSS
VLDNVPLVAATMGMYTLPLDDTLWKLIAYAGTGGSILIIGSAAGVAFMGLEKVDFLWYF
KRHGWIALASYFGGLFSYFVLESLMFFI KRIEWIALASYFGGLFSYFVLESLNFFI

1167027 1168898

CPHE 016

1167027 1168898

CTBSB hypothetical protein

KREW-MKKGKLGAIVFGLLFTSSVAGFSKDLTKDNAYQDLNVIEHLISLKYAPLPWKELL

FGWDLSQOTQQARLQLVLEEKPTTNYCQKVLSNYVRSLNDYHAGITFYRTESAYIPYYLK

LSEDGHVFVVDVQTSQGDIYLGDEILEVDGMGIREAIESLRFGRGSATDYSAAVRSITSR
SAARGDAVPSGIAHLKLRRSCLIRSTPVRWRYTPEHIGDFSLVAPLIPEHKPQLPTOSC

VLFRSGVNSQSSSSSLFSSYMVPYFWEELRVQNKQRFDSNHHIGSRNGFLFTFGPILWEQ

DKGRYRSYIFKAKDSQGNPHRIGFLRISSYVWTDLEGLEEDHKDSPWELFGEIIDHLEKE

TDĀLTIODTHNPGGSVFLYSLLSMLTDHPLDTPKHRMIFTQDEVSSALHWQDLLEDVFT

DEDALVAVLGETMEGYCMDHAVASLQNFSQSVLSSWVSGDINLSKPMPLLGFAQVRPHPK

HQYTKPLFMLIDEDDFSCGDLAPAILKDNGRATLIGKPTAGAGGFVFQVTFPNRSGIKGL

SLTBSLAVRKDGEFIENLGVAPHIDLGFTSRDLQTSRFTDYVEAVKTIVLTSLSENAKKS

EEÖTSPQETPEVIRVSYPTTTSAS EEOTSPOETPEVIRVSYPTTTSAS

CPT_017 1168997 1169935 19th Metalloprotease VIIMRKLILCNPRGFCSGVVRAIQVVEVALEKWGAPIYVKHEIVHNRHVVNALRAKGAIF VEELVDVPEGERVIYSAHGIPPSVRAEAKARKLIDIDATCGLVTKVHSAAKLYASKGYKI LIIGHKKHVEVIGIVGEVPEHITVVEKVADVEALPFSSDFUFTYITOTTLSLDDVGFISS ALLKRYPSIITLPSSSICVATTNROKALRSVLSRVNYVVVVGDVNSSNSNRLREVYLRRG VPADLINNPEDIDTNIVNHSGDIAMTAGASTPEDVVQACIRKLSSLIPGLQVEND DVVFQLPKELRCS

1169895 1170629

NO robust homolog present in Genebank/EMBL as of 11/7/98
RMSYFNYOKNSVVLRSLGLLAKFFSRLLYRVFFSFREGIYLFSSLYLKYPRYFFYDLGKY
VYSLRHCPYAKLGRLPGASLLKEGNVYGETPWSVLAKICQAFDITSODILYDLGCGLGKV
CFWFSHVVRCQVIGIDNQPHFIRFSSNMHRKLSSGFALFDTEEFKNVVLSASYVYFYGS
SFSRRLLNEIILKLSEMAPGSVVISISFPLDSFSRGKECFFTEKSCSVR*PWGKTIAYKN

1172146 1170638

CPH_1019 1172146 1170638
CTB60 hypothetical protein
IHRRNIMTVSYQSISTPPPEGEFDIFVDCNATEEAV/AAEVQVALBAGEQYAMLRATSEL
CFGILITOGECALTOALPPREREJQEEQEYLVNGILMRSTSLENLKYGOSQGTSLASHBNP
LAQQCTGCSCATTOALPPREREJQEEQEYLVNGILMRSTSLENLKYGOSQGTSLASHBNP
LAQQCTGCSCATTOALTSSSFPFFSCKAPEGGS/DKTFTVSYGTFKAQECQEASASQ
GOAQFIVRSYGGSTIKEHGAKEKVSQSTKGAETQKHTQTKGDATLSPMSLYSTLHKEVPQ ALICTROQKIDEEHROURQGEGYEQEGGEGERKKTPWCTVSJAQVTGSSNGVYESYTPI IPDP1VEFALICEQLISVLAGKRVTNLDVLRICTELMFLMLKSJAANDTMTRLEERELMERE AHELAAGYGRQAKYARWIGITATATLGILGATATMVGEIGGGGTLGFVQRIGGPFKDATAK TELKO DIKALLOPITEVADKAHEPREDAABKARAKAEAKKEALKWINDEALLI TEEAKDIM KOMUNELLNILOPEHDAARSLYO

117 604 1172150

CTHELTORO 117 0-94 117.2150 CTBGT Typochet igit profesia CGBGTHGSMIGGAGEVILAQDEY LEDARBOGEBYEK DYG ETVARQBAÇKOLPKEFTOK FOSOCHGBGETHEKTELLATERER LERGOGEBGGEBGGINTSOAQTBGFWNLFSOKNSTBA GKALMOBITMEKGPEKTGEKALDKNIGSKOPSIGKNYDTERLAGHLKLEGTYDGLYSOSL

DSEQUELLQSRREERSETY SEKNIETHVQIKDLUKDLFSQDQDSNCKQKKSPFQ DTGRKNRIAKAAQAVPVIP FTLSYLLTKQGILSDFSSYGCHKDSVESTQRELDA LHEKRIETIKVSIEKEKRERLAGSLSDIIGWLAPFVSIGIGIVAILSGGGIFAFAGFFAG LISLVIKCLEKLKFWDWLEKHLPIKNEELKRKIITIIGWVVYLTFVILSICTLKVENLGF SPIIEGAIGIPAIESTMAALRCAILESOAEIYKLKGKLTKIQLDIELKSFDRDDHYER SQELLDNMESSFEALSRILMYMRELDQV/LHSLRG

1174270 /1173698

Primiturw (Nal. 1986) Rollinghor Newschild Community PRIDAL INFORMATION LEGYLEGY/FORMIALD TYOKIFKISSEDLEKYYKEUYHAYLOKDYAKUITVFKWLVFFNPFVSKFWFSLGASLHMS EQYSQALHAYGVTAVLROKDPYPHYYAYICYTLTNEHEEAEKALEMAWVRAQHKPLYNEL KEEILDIRKHK

1175/09 1174216

CPn_1022 1175/09 1174216
CT863 hypothetical protein
FSFFFYALKLQIMMMPVPSAVPSANITLKEDSSTVSTASGILKTATGEVLVSCTALEGSS
STDALISLALGGILATQGE/LLQSTNVHQLLFLPPEVVELE:QUVDLLVQLEHAETITS
EPQETGTGSRSEGTLPQQSSKQSALSPRSLKPEISDSKQQALGTPKDSAVRKHSEAPS
PETQARASLSQASSSSGSLPPQESAPERTLLEQQKASSFSPLSGFSAEKQKEALITTSKS
HELYKERQDRQQREQHDYKHDQEEDAESKKKKKKRGIGVFAVAEEPGENLDIAALIFSD
QMRPPAEETSKKETTFKKLESPMSVFSFFIESKNPLSVGSSIHGPIGTFKVENVTFERM
KLMARILGQAEAEANELYMRVKQRTDDVDTLTVLISKINNEKKDIDWSENEEMKALLNRA
KEIGVTIDKEKYTMTEEEKRLLKENVQMRKENMEKITQMERTDMQRHLQEISQCHQARSN
VLKLLKELMDTFINNIRP VLKLLKELMDTFIYNVRP

CPn_1023 / 1176008 1176331
No robust homolog present in Genebank/EMBL as of 11/7/98
GLDFLEIFIMKKY/TLSIIFFATYCASELSAVTVVAVPLSEAPGKIQVRPVVGLQFQEEQ GSVPYSFYYPYD/GYYYPETYGYTKNTGQESRECYTRFEDGT I FYECD

CPn 1024 1177317 1176334

CPn_1024 1177317 1176334

xerD-Integrase/recombinase
IFFFPWFSL/SLKIAPLPILKLHSLASMTMPSTOFHTTILEOFSLFLSVDRGLCQQSIAA
YRQDISSFL/TISAISSPODISQNSVYIFAEELYRRKEAETTLARRLIALKVFFLFLKDQQ
LLPYPPIIEHPKIWKRLPSVLTPQEVDALLAVPLQMENPRHLAFRDTAILHTLYSTGVR
VSELCDL/BLGHVSDDC IRVTCKGSKTRLVPLGSRAREAIDAYLCPFRDQYQKKNPHEDHL
FLSTRGHKLERSCVWRRIHNYAKQVTSKPVSPHSLRHAFATHLLDNKADLRVIQEMLGHA
RIASTE/YTHVAADSLIEKFLAHHPRNL

CPn_1025 1177266 1178879
pgi-Glucose-6-P Isomerase
GAEQFSSYREKTMERKRFIDCDSTKILQELALNPLDLTAPGVLSAERIKKFSLLGGGFTF SPATERLIDAILAALISLAEERGLIESMLAMOGQVVNYIEGFSEMPALHTATRAWT DSÉFTGEAEDIAVRSKVEAQRLKDFLTKVRSOFTTIVOIGIGGSELGFKALYRALRAYCP TÜRKTVHFISNIDPÜNGAEVLÖTIGKALLVVVVSKSGTTIETAVNEAFFADYFAKKGLSF VEHFIAVTCEGSPMEDTGKYLEVFHLWESIGGRFSSTSMVGGVVLGFAYGFEVFLOLLO SAMO IALOPNARENLEMLSAL ISTWINNFLOYPTEAV I PYSSGLEFF PAHLOCCHES
(MCKSTAODGRAVGFSTSPVIMGEFGTINGOMSFFOCLINGGTD I FVEF IGFEKSOKGED IS
FOOTISSOKLFANMI AQAI TALACGSENTIPENKINFDORNEPSSVI VSSQLINPYSLGELLSYY
ENKIVFOGLCWG INSFDQEGVSLGKALANRVLELLEGADASNFPEAASLLTLEFNIKFY

1178961 1179137

CSFGFGKICEDRMFFIAVRSRGFLDIHGILAARKGKQVVKSTAGAWIGSRGAVFYSLVS

1179172 1180755 No robust homolog present in Genebank/EMBL as of 11/7/98 NMPGSVSSPPLSPVIVRERVPSSSGSDLIQPHAVLKISILIFALVTILGIVLVVLSSALG

MMGSVSSPPLSPVIVRERVPSSSGSDLIQPHAVLKISILIFALVTIGIVLVVLSSALG
ALPSLVLTVSGCIAIAVGLIGLGILVTRLILSTIRKVDAMGYDAAVKEQYLSRIRELES
ENREIRDRNRAVEDQCAHLSEENKOLRDPEYLHGMTERLIASLEIENQALVAENILLKDW
NASLSRDFRAYKQKFPLGALEPWKEDIACIMEONLFLKPECIAMVKSLPLETQRLFLYPK
GFQSLVNRFAPRSRFFQTPKYEYNSRNENEDGKVAAVCARLKKEFFSAVLGACSYELLGG
ICERAVALKETLPLPEAVYDTLVQEFPNLLTAESLWKEWCFYSYPYLRPYLSVDYCKRLF
VQLFEELCLKLFTTGSPEDQALVRLFSYYRNHIPAVLASFGLPPPETGGSVFVLLPKQEN
LLWSQIEVLATRYLKDTFVRNSEWTGSFEMMFSYNEMCKEISEGRIRFAEDYETRHSEEF
PDSPLISFEGEFFEFI DROSSFFFVSVI FEDDTI NYDEMMANJEDJRUKCEI PPSPLSEEGEGEFLPPCSEEEVSVLERPDLDVDSMWVWHPPVPKGPL

1180995 1181999

mdhC-Malate Dehyrogenase FFLKGVRMAFKEVVRVAVTGGKGQIAYNFLFALAHGDVFGVDRGVDLRIYDVPGTERALS GVMMELDDGAYPLLHRLRVTTSLNDAFDGIDAAFLIGAVPRGPGMERGDLLKQNQQIFSL GGALNTAAKRDAKIFVVGNPVNTNCWIAMKHAPRLHRKNFHAMLRLDQNRHHSNLAHRA EVPLEEVSRVVIKGNHSAKQVPDFTDGARISKKPAAEVIGDROMLENILVHSVQNRGSAVI EARGKSSAASASRALAEARRSIFCPKSDEWFSSGVCSDHNPYGIPEDLIFGFPCRMLPSG DYETTPGLPWEPFTRNKTQTSLDETAQEKASVSSL

CPn_1029 1181987 1182844

No robust homolog present in Genebank/EMBL as of 11.7/98

RVFVISTMLWGVSHROSFDELSQNAFKNIFNKQRFCFIFCSLCCFGFVFALFLKLCSRLA

PEISLSTLGLGAFFCAFSVICASAIIVGFLLHKESGGETSKLCCAIKNTWSSLWLSLLVS

MPFFIANVAVVTVAMLSSFLCSLPWVGKLFHTVLIFIPVLSATALILLIEJSFSCLFFCI

PVLHNQESIDYRKLLECFRGNILROFIGVVIALVPLALCSWLALDSFYLMTHLVEIADIH

TWSFLAQMFVLIVPIALILTPAVSFFFNFSFSFYLAKQEEEKALVK

1183901 1182843

CPH_1030 1183901 1182843
predicted D-amino acid dehyrogenase
FKVHFMRIAVLGACYACLSVTWHLLLHSOCTATIOLFDFIFLGEGASGMSSGLLHAFTGK
KALKPPLADQGINNTHALITEASKALNYPIVISGGILRPAIDEDQAQLFTERVEEFPKEV
EMWEKARCEISIPSMVIPPNLGALFIKSSVTULNBLUFYGGLADANKKLTTQFYDELIEDL
ADIEEFYCHTIVTTOANASILPELKEMPVNKVYGLLEIJADANKLTTQFYDELIEDL
NTQKNTCILGATFEHNQPEETPDPAIAYQEIMPPVLSLFFGLKDAQVLHCYALMRSSKS
RLPVISRIREKLWFIXSLGSKGLLYHGITTSDMLAGAVLRKSTAYIAKEFLFTI

CPn 1931 1185567 1184098

THANDALLY THE PERSON STRANDS THAN THAN SALE WAS TO SELECT A THE STRAND STRANDS TO SELECT A SE

OLITAVLOOWLGWTTIVAETPFGAAKNOTFPEIFTIE VSLYITSSVMOLAML VYFOONAWITMLSITGVMVLPAYLAGAAFLFKLSKSK SLWLIYAGGLKYLFMALVELALGIPFYIDAGKKKKNAK TKAPLAMITG LLGVVY

ALGELNFENAEPAKVN

1187656 1186187

CPn_1033 1187556 1186187
CT372 hypothetical protein
NNKKKDYSGEFLTTDTVDSIAFLPSEENFCYIKTILFFRVKKKHYAFFYGEFMISFRFLL
LSGLCALGISSYAETPKETTGHYHRYKARIQKKHPESIKESAPSETPHHNSLLSPVTNIF
CSHWKKGISVSNLLTSVEKATNTQISLDFSILPOWFFYPHKALGOTQALEIPSWOFFYFSP
STTMTLVSDPTAGGGIVDFSYTLIHWQTNGVDANQAAGTASSMDVSNRENNLAQLIFS
OTFPGDFLTLAIGQYSLYAIDGTLYDNDQYSGFISYALSQNASATYSLGSTGAYLQFTPN
SEIKVQLGFQDSYNIDGTNFSIYNLTKSKYNFYGYASWTPKPSCGDGQYSVLLYSTRKVP eonsovtgwslnaaohiheklylfgringatgtalpinrsyvlglvsenplnrhsodllg igfatnkvnakaisnvnklrryesvmeafatigfgpyisltpdfolyihpalrperrtso VYGLRANLSL

CPn_1034 1188589 1187732

Predicted OMP (CT371) (leader (18) peptide)

KTSWOKYKKYLSYSILVOKIARYVMKTWLFFFFLFSCSSFYASCRYAEVRSIHEVAGDIL

YDEENFWLIIDLDDTLLQGEALSHSIWKSKAIQGLQKQGTPEQEAWEAVVPFWIEIQEM

GTVQPIESATFLLIEKIQKQGKTFFYYTERPKTAKDLTLKQLHMLNVSLEDTAPQPQAPL

PKNLLYTSGILFSGDYHKGPGLDLFLEICTPLPAKIIYIDNQKENVLRIGDLCQKYGIAY

FOLTYKAQELHPPIYFDNIAQVQYNYSKKLLSNEAAALLLRHQMHE

CPn_10 19081 1188570
aroE-Shikimate 5-Dehyrogenase
\text{VVOLDEMYPIVHLOIWRFSMIYYGVSVMICATVSGPSFCEAKQILKSLHLVDIIELRLD}
\text{VINEDDOELHTLITTAONPILTFRQHKEMSTALWIQKLYSLAKLEPKWMDIDVSLPKTA}
\text{LOTTRKSHPKIKLILSYHTDKNEDLDAIYNEMLATPAEIYKIVLSPENSSEALNYIKKAR
\text{LLHKPSTVLCMGTHGLPSRVLSPLISNAMNYAAGISAPQVAPQOPKLEELLSYMYSKLSE
\text{KSHIYGLIGDPVDRSISHLSKNFLLSKLSLNATYIKFPVTIGEVVTFFSAIRDLPFSGLS
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL VAGNUSINDEN
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANYICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISETURIEMPANTICATTSCHAMIL
\text{LTMLVENTYLEMENDLD DASAGISTURIEMPANTICATTSCHAMIL
\ RSHIYGLIGDEVORSISHLENHILLSKLSLAMIYINFPYTIGEVYPTSAINDLFFSALIDDLFSSLS
VTMPLKTIFFPHVDALDASAQLESINTLVFRNOKILGYNTDOBGYAKLIKONISVANK
HIAIVGAIGAMAIAATLANGGANLHIFNRTLSSAAALATCCKGKAYPLGSLENFKTIDI
IEKGIPPHVDFWFFPFIVMDINTKPHPSPYLERAQKHGSLIIHGYEMFIEQALLQFALW
TRALMYESCDSFRNYVKNFMAKV

TRUEM FESCDSFRNYVKNFMAKV

CPT=1036 1191180 1189984

ards Dehyroquinate Synthase
GYDKFCSCRSCIIPTMLOTMNSETIITTPHVVKLISNFFQKKLFSSISTAYPLVIITDVS
VÖMLLGPILDHIKNLGYQVIVLTFPFGEPNKTWETFISLQYQLVDONISPKSSIIGIGG
GTWEMTGFLAATYCRGLPLYLIPTTITAM/DTSIGGKNGINLRGIKNRLGTFYLPKEVW
MCPWGFLSTLPREEWYHGIAEAIKHGFIADAYLWEFLNSHSKMLFSSSQILHEFIKRNCQI
KAALVAEDPYDRSLRKILNFGHSIAHAIETLAKGTVNHGQAVSVGMIETRISLAEGVMK
TPQLIDQLERLLKRFNLPSTLKDLQSIVPEHLNNSLYSPENIIYTLGYDKKNLSQHELKM
IMERICRAAPFNGTYCASPNMEILYDILWSECHVMRHC

1192286 1191123

CPREMUS/ 1192250 1191123
arco-Chorismate Synthase
LHFSRGSRRSFLEELLRTSVSRSHYLVKVMKNSFGSLFSFTTWGESHGPSIGVVIDGCPA
GLELHESDFVPAMKRRRPGNPGTSSRKENDIVQILSGVYKGKTTGTPLSLQILNTDVDSS
PYENSERLYRPGHSQYTYEKKFGIVDPNGGRSSARETACRVAAGVVAEKFLANQNIFTL AYLSSLOSITIENTYLKISPELIKKITTSPTYSPLPNEKIGELITSLHDDSDSLGGVISFI TSPTHDFUGEPLFGKVHALLASALMSIPAAKGFEIGKGFASAQMRGSQYTDPFVMEGENI TLESENGCOTLGGITIGVPTEGRIAFKPTSSIKRPCATVTKTKKETTYRTPQTGRHDPCV AIRAVPVVEAMINLVLADLVLYQRGKKL

CPĒ_1038 1192750 1192199
aroll-Shikimate Kinase II
wKLELRNVMTIILCGLPTSGKSSLGKALAKFLNLPFYDLDDLIVSNYSSALYSSSAFYK
AYĞDĞKFSECEARILETLPPEDALISLGGGTLMYEASYRAIGTRGALVFLSVELPLIVER LEKRGLPERLKEAMKTKPLSEILTERIDRMKEIADYIFPVDHVDHSSKSSLEQASQDLIT

CPn_1039 1194011 1192665
aroA-Phosphoshikimate Vinyltransferase
vCFTMLTYKVSPSSVYGNAFIPSSKSHTLRAILWASVAEGKSIIYNYLDSPDTEAMICAC
KQMGASIKKFPOILEIVGNPLAIFFKYTLIDAGNSGIVLRFWTALACVFSKIITVTGSSQ
LQRRPMAPLLQALRNFGASFHFSSDKSVLPFTMSGPLRSAYSDVEGSSOGPASALAVACS
LAEGPCSFTIIEPKERPWFDLSLWWLEKLHLPYSCSDTTYSFPGSSHPQGFSYHVTGDFS
SAAFIAAAALLSKSLOPIRLRNLDILDIGGDKIFFSLMONLGASIQYDNYEILVFPSSFS
GGSIDMGCIDALPILTVLCGFADSSHLVNARSAKDKESDRILAITEBLGKKGACIOPT
HDGLLVNPSFLYGAVLDSHDDHRIAMALTIAALYASGDSRIHNTACVFKTFPNFVQTLNI
MEARIEECHDNYSMWSTHKRKVFARESFG

CPR_1040 1194876 1194073
NO robust homolog present in Genebank/EMBL as of 11/7/98
RPSOSLFLRTWSPSSFREHTVCAAPLLYPRRESPBYLFSPTGC/MSTTTYKHFIHTASR
WEPVLKEIVASNYWHAQWINTLSFLENSGAKKISASEHPTEVKEVLKHAAEERRHGHYL
KTOISRISETSLPDYTSKNLLGGLITKYYLHLLDLRTCRVLENEVSLSGOTLKTTAYTLV
TYATELRASELYPLYHDILKEAGSKITVKSITLEEQGHLQEMERELKDLPHGEELLGYAC
OFEGELCLDFVERLEGMIFDPSSTFTKF

CPm_1041 119630 - 1194726 *bioA -Adenosylmethionine -8 - Amino -7 - Oxongmanoate

Aminot construction

DERRIE FPUHERKES MCVFFLVMFLDESEPS LSRKYCHLENGRSNALTKPOYLRYOCK WYSRIPALANSPAQASYMEDGG ERMELDESEE ISK ZAJALERGENALTKEGYTGYGGK WYSRIPALANSPAQASYMEDGG ERMENBENRIKLEGYMEGOMOGOSGGNGC UMPET GEALDTPER IVEGE AYUZAGSGTRYLDA LOSWACALIGHGHEY ITKKLCEGAGKLEHV FRANFTHEPALELVSKLAPILLEGLERFFEDONG TØ IE IAKKLAVGYYNGIKAKSHEV GLENAZIG TEFAMELATGITTIVPFÜDLEFLEST ZAPYZYKERELA IAQAATVESEGNI AAF ZEFIALA VARMIMYNEG IKKELLKLARHYGYLG IADELLTGEGRTGPLFASETTDI PROFICES WAS A PROPER A DEVET OF THE PROPERTY VICTVIALDYPARATOYFSQYRDHLN

REFLERGVLLRPLGMTLYV DEEDLRY LYCHLODALCLOPO

1126629 1195234 bioD-dethiobiotin synthetase

*Diod-decriodictin synthelasy MRSPFTYFRANFFMORIIIVGIDTGVGK*IVSAILARALNAEYWKPIQAGNLENSDSNIV HELSGAYCHPEAYRLHKPLSPHKAAQIBNVSIEESHICAPKTTSNLIIETSGGFLSPCTS KRLOGDVFSSWSCSWILVSQAYLGSINHTCLTVEAMRSRNLNILGHVVNGYPEDEEHWLT QEIKLPIIGTLAKEKEITKTIISCYAEOWKEVWTSNHOGIQGVSGTPSLNLH

DIOF_2-OXONONANOEC SYNTHASE_2
PMLCOOFLIEALARRKSKHTYRSISLNSHLIDFTSNDYLGFASSPELRKEYITKLHAIES
LGATGSRLITCHSOLCORIEED(MAYHNFESCLIFNTGYTANLGLLYALATDCORILHDL
YHASIYUGIRLSKAOSFPFHYMDIUNHLEKRLASSHLGRIFVCVESVYSLHGSVAPLQAI
SELCERYSAYLIVDEAHAVGVFGDGEGLVSALGLODKVLATVYTFGKALGTHGAAIAGS
SILKDYLINFGRFFIYTTAGYPHALTAIELAYEHNORAFNQREHLSALIHHFREKAONLG
LQLMKDNTTTFIOSICVSGSMRARQAALQIQNSGYDVRFIVSPTVKQREELLRICLHAFN
TKNEIDHLLHTLEQIFLCNYSSL

1198700 1197699

**DioB-Biotin Synthase
AKHMREETVSWSLEDIREIYHTPVFELIHKANAILRSNFLHSELOTCYLISIKTOGCVED
CAYCAQSSRYHTHTYFEMMKIVDVVERAKRAVELGATRVCLGAAWRNAKDDRYFDRVLA
WYKSITDLGAEVCCALGMLSEEQAKKLYDAGLYAYNHNLDSSPEFYETI ITTRSYEDRLN
TLDVVNKSGISTCCGGIVGMGESEEDRIKLLHVLATRDHIPESVPVNLLWPIDGTPLQDQ
PPISFWEVLRTIAYARVVFPRSMVRLAAGRAFLTVEQQTLCFLAGANSIFYGDKLLTVEN NDIDEDAEMIKLLGLIPRPSFGIERGNPCYANNS

CPn_1045 1199602 1198901

CM_1045

**Conserved hypothetical bacterial membrane protein
GTLPMNTSHRKTLVFSYLSSTFTLLLVLSNLVLSSKLIPTTFFNFIIPOGLILYPLTFLI
SDVVNEIFOFKKARVMIFSAFIANLLASSIVOIFMFFPVASPEMQTAMACLFDLSPLRFL
ASLLAFIVYOQLDIVLYTFFKNRTFNSSLWLRSNOSTWISQIPDTFIVDTCILYFGMGLS
FPQTLNIMFYSYIYKITFCVLTTPLFYLAVNTIRKFLGMPSTKIANTVPLINQP

1200675 1199590

TTYPTOPDAM HYTOXYlase

VHYCERTLDPKYILKIALKLRQSLSLFFQNSQSLQRAYSTPYSYYRIILQKENKEKQALA
RHKC/SILEFFKNLLFVHLLSLSKNQREDCSTDMAVVSTPFFNRNLWYRLLSSRFSLWKS
YCPHFFLDYLEAFGLLSDFLDHQAVIKFFELETHFSYYPVSGFVAPHQYLSLLQDRYFPI
ASVMRTLDKDNFSLTEDLIHDLLGRVPMLLHPSFSEFFINNGRLFTKVIEKVQALPSKKQ
RIJTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPQELGHAFIDNVRVLPLEL DOTIERPFINISTPOETLFSIRHFDELVELTSKLEWMLDQGLLESIPLYNQEKYLSGFEVL

CPn_1047 1200537 1201343
dapB-Dihydrodipicolinate Reductase
FGSRNMGSSMHVGVIGCSGRTGKVIVSALEQSSEYTLGPGFSRSSALTLFQVIAHNDVLV
DFSHPLLTKEVVAHLLISFKPLIIGTTGFPGKCKEAHDSLEELTHIVPVVVCPNASLGAY -ekkrlvmilsqicnpofdirirethhrykkdslsgtaqdlldtiqqvkqedwgeeyevgo rdgskktievqssrvgdipgehevafissgeqilvrhtvfsrnvfgrgilsildwlktln POGLYSLGDTLELVLRNEHCLLKKTTDH

1048 1201588 1202604

asd-Aspartate Dehydrogenase
DGERKGMRIAVLGVTGLVGQKFVALLHKWYRDWVIAEVVASNSKYGQSYGDACIWQEPIG GERNAMKIAVISVIGLVOGUT VALLHAMIKUMVIAEVVASNSKIGGSIGDALIMGEIG
PMPEMYRDLPIRKIEEFVQSDIVVSFLPSSAESMEAYCLSQGKVVFSNASTYRMISSVPII
IPEVNSDHFQLLEEQPYPGKIITSPNCCVSGITLALAPLRKFSLDHVHIVTLQSASGAGY
PGVPSLDLLANTVPHIVGEEEKILRETVKILGSSKQPLPCKLSVTVHRVPVAYGHTLSLH
VTFSKDVDLDEILYSVQEKNKEFPHTYQLVDNPWSPQARKHLSHDDMRVHLGPITYGGDF
RTIKMVLIHNLVRGAAGTLLASMENYFFDYLKREMCLR

1202586 1203914

CPH_1049 1202586 1203914

lysC-Aspartokinase III

EGNVSKIVYKFGGTSLATAENICLVCDIICKDKPSFVVVSAIAGVTDLLVDFCSSSLRER
EEVLRKIEGKHEEIVKNLAIPFPVSTWTSRLLPYLQHLEISDLDFARILSLGEDISASLV
RAVCSTRGWDLGFLEARSVILIDDDSYRRASPNLDLMKAHWHQLELNQPSYIIQGFIGSNG
LGETVLLGGGGSDYSATLIAELARATEVRIYTDVNGIYTMDPKVISDAQRIPELSFEENQ
NLASFGAKVLYPPMLFPCMRAGIPIFVTSTFDPEKGGTWVYAVDKSVSYEPRIKALSLSQ
YQSFCSVDYYLGCGGLEEILGILESHGIDPELMIAQNNVVGFWDDDIISQEAQEHLVD VLSLSSVTRLHHSVALITMIGDNLSSPKVVSTITEKLRGFQGPVFCFCQSSMALSFVVAS ELAEGIIEELHNDYVKQKAIVAT

CPn_1050 1203884 1204798
dapA-Dihydrodipicolinate Synthase
LCKTKSYSRHVGRIMHLLTATVTPFFPNGTIDFASLERLLSFQDAVGNGVVLLGSTGEGL
SLTKKEKQALICFACDLQLKVPLFVGTSGTLLEEVLDWIHFCNDLPISGFLMTTPIYTKP
KLCCQTLWFEAVLNAAKHPAILYNIPSRAATPLYLDTVKALAHHPOFUSIKDSGSVEEF
OSYKSIAPHIQLYCGDDVWSEMAACGAHGLISVLSNAWPEEAREVVLNPOEQDYRSLWM
ETCRWYYTTTNPIGIKAILAYKKAITHAQLRLPLSIEDFDLENVSPAVESMLAWPKLRTS VESYS

CPn_1051 1204956 1205270
No robust homolog present in Genebank/EMBL as of 11/7/98
FFMTPKSIQQLHLIKTIDPVRKISPVTTKKSSFFRQSLLRFLELFWMFLYCIRSIRPHCV HIATFICRGLILFLTTLFLSMICILHFITLPWICKEDPRIIRKNK

CPn 1052 1205402 1206169

CPD_1052

LIGS402

LIGS403

NO rOBUST homolog present in Genebank/EMBL as of 11/7/98

FFIGKMKYNSREKIK:ALRICSSYCITVERNNESLSCYDKIFYSLSCYVERKENSIGRCR

SPCFFRCKKTEVETKEVKIKEE IREDIEGNDFVKVAESE PERMAALE:LSSGOSIGNLCA

GONELOSGMESRESKETWEST IFFROSTTOABETSBEFFRETACHIA-HEROTIA-HEROKLAGSYEL

GVTAGLLGGREKOVEDSHRTRATSSIILSWYTRPLSCTYFYIVGKARPLLEFFRETSD VRRDLKKKERLEFCKD

(20._10%) 1206138 1206701 NO TODIOT bomolog present in Genebank/EMBL as of 1177798 KKOODEMPAKTDINHLYLTYCLDD/IVACH-LETTXLLTMYGEKASHEVLYYSKFRETIOGEP GELATS/NDTYYSTIVSLAFGLGCEFFFDE/GRIDDHIDHIJAAK GAVLIBETREAKETPG GOKDYAFFGLTARESIMISEKLAMTFG/VSEVIQHCYSGCTFY/FKTNIKEQYRHISHMYGF

CPn_1054 1207010 1207465

No robust homolog present in Genebank/B. as of 11/7/98

SRWHRFE IMOVILSPQLPPPPGHSVGSISSPSKLRVLAITFLVFCMLLLISGALFLTLGI
PGLSAAISFGLGIGLSALGGVIMISGLLCLLVKREIPTVRPEEIPEGVSLAPSEEPALQA
AQKTLAQLPKELDQLDTDIQEVFACLRKLKDSKYESRSFLNDAKKELRVFDPVVEDTLSE
FELDRIVAQEGWDLINE INGGRSLMMTAESESLDLFHYSKLGYLPSGDVRGEGLKKSA
KEIVAPIMISHCEHKVAVAFDRNGVAMAEKAFAKALGALESVYRSLTGSYRDKFLESE
FOR DEM BEITPERBUDAE PLABER BERBERFFFANDAVEED FPROINDLOUGHT
VVROMYDQEFOKAGERLEKLHALYPEVSVSIRENKIQETRSNLEKAYEAIENTRCCVRE
QEDVWKEEKKRAEFRERGNKILSPEELESSLEGOFDHGLKNFSEKLMELEGHILKLOKKA
TAEVENKILSDAESRLEIVFEDVKEMPCRIEEIEKTLRMAELPLLPTKKAFEKACSQYNS
CAEMLEKVKPYCKESLAYVTSKERLVSLDEDLRRAYTECQKFFQGDSSLESEVRACREQL
GRINMTMSERLREGVQACNKMLKAGLSEEDKVLKEEEYWLYREERKNKEKRLVGTKIVAT
QGRVAAFESIEVPEIPEAPEEKPSLLDKARSLFTREDHT

CPn_1055 1209583 1210521
No robust homolog present in Genebank/EMBL as of 11/7/98
CKYLYHHSYPPPDHSVGAFFCLSKFRVLAITFLUCGVLFLISGALFETIGISGLSAAIS
FGLGIGLSALGGVLWVSGLLCLLAKREVPTVRPEEIPEGVSVAPSEEPALQATOKTLAQL
PKELDQLDRVIGEVSCLGKLKDLRECEDGGLLKDAKEKLQVPDFVWKDMMTEFVELQQIM
DQEGWYLKCLIQEMRDIGSTLFMSQVSLFKLMEWLGYLPSGDVRGERLKKSAREVVDRFM
RRICDTRKVAMTFDRNAYGVAKTAFEKAFGALETCVYKSMTESYREAFCEYKKTKILRDE
EKILRICYLELRR

CPn_1056 1210482 1211228

NO robust homolog present in Genebank/EMBL as of 11/7/98
GEDIKDMLSRVEELEMMLRVIELPLLPIKQALEKAFVQYNSYKAKLTKVEPCFRESPAYI
TSEERLQSLOOTLERAYKEYOKRFQEPSRLESEVSGCREHLREQVKQFFTGGLDLIKEEL
IFVSDVLFRKMVSCLVSTVHVPFMEFYYEYFELHRLRLRAQMMANAEIYSKVRKAFPEML
KETLEKAKAPREEEYWLLCEERKSKEKRLILNKIEAAQQRVKDLEPPPIKETGKQKRKKE
YSFFIRLKS

CPD_1058 1213742 1214836
CT835 hypothetical protein
EVMRLYGTLEGIVLVSTGC:IFLOMHGGYAAEVPVTSSGYENLLESKEQDPSGLAIHDRIL
FKVDEENVYTALDVIHKLNLLFYNSYPHLIDSFPARSGYYTAMWPVVLESVIDEFLMVAD
AKARRIATDPTAVNQEIEEMFGRDLSPLYAHFEMSPNDIFNVIDRFLTAQRVMGMVASK
VMLKVTFGKIREYYRKLEEEASRKVIWKYRVLTIKANTESLASQLADKVRARLNEAKTWD
KDRLTALVISGGGGLVGSEFSRENSELSGSHKQELDLIGYPKELGGLPKAHKSGYKLYBL
LLDKTSGSIEPLDVMESKIKQHLFALEAESVEKQYKDRLRKRYGYDASMIAKLLSEEABP

CPB-1059

1214848

1215678

kgsA-Dimethyladenosine Transferase
VTRSSPAQLSRFLSEIQNKPKKSLSQNFLVDQNIVKKIVATSEVIPQDWVLEIGPSFGAL

TEELIAAGAQVIAIEKDPMFAPSLEELPIRLEIIDACKYPLDQLQEYKTLGKGRWANLP
YHİTTPLLTKLFLEAPDFWKTVTVMVQDEVARRIVAQPGGRDYGSLTIFLQFFADIHYAF
KVSRSCFYPKPQVQSAVIHMKVKETLPLSDEEIPVFFTLTRTAFQQRRKVLAMYLKGLYP
KEQTEQALKELGLLLUVVRPEVLSLNDYLALFHKMQAG

CHITO60 1217694 1215727

dxs/tkt-Transketolase
YKRPLYIHITKVMTSSSCPLLDLILSPADLKKLSISQLPGLAEEIRYRIISVLSGTGGHL
SSNLGIVELTIALHYVFSSPKDKFIFDVGHQTYPHKLLTGRNNEGFDHIANDNGLSGFTN
PTESDHDLFFSGHAGTALSLALGMAGTTFLESRTHVIPILGDAAFSCGUTLEALMISTD
LSKFVVILDNOMMSISKNWGAMSRIFSRWLHHPATNKLTKQVEKULAMIPHYGOSLAKHS
RRLSQCVKNLFCPTPLFEQFGLAYVGPIDGHNVKKLIPILQSVRNLPFPILVHVCTTKGK
GLOQAQNNPAKYHGVRANFNKRESAKHLPAIRPKFSFPDIFGGTLC£LGEVSSRLHVVTP
AMSIGSRLEGFKQKFPEFFFDVGIAGEHATVTFSAGIAKAGNPVICGTVSTFLHRALDNVF
HDVCMQDLPVIFAIDRAGLAYGDGRSHHGIYDMSFLRAHPQMIIGPRSQWFQOLLYSS
LHWSSPSAIRYPNIPAPHGDPLTGPNNFLRSRONAETLSQGEDVLIIALGTLCFTALSIK
HQLLAYGISATVVDPIFIKPFDNDLFSLLMSHSKVITIEEHSIRGGLASFFNNFVATFN
FKVDILNFAIPDTFLSHGSKEALTKSIGLDESSMTNRILTHFNFRSKKQTVGDVRV

CPn_1061 1217932 1217666
CT330 hypothetical protein
FGSLMVEIHHKDPSLKKLFALQQSLETLNSLSDIVATYEAMFSLIYEGLNKALRKDQLCY
LLSWNSKGELLKSPSGDPTVQTFFIHPHH

CPn_1062 1219835 1219159
xse4-Exodoxyribonuclease VII
RGFPVMIDPPOAVASTER IKTLLESNFCQIIVKGEUSNVOLQPOGHLYFGIKDSQAFLN
GAFPHEDIKYYDERPKOGDAVI THOKLAVYAPROQYJIVAIIALVYAGEODLLOKFEETKR
HLTAEGYFATEKKKPLPFAPQCIGVITOPTGAVIO LIEVLORRARNYKILVYPVTVQON
DAAHEIDKA LEVMBAENLADVLITARGOGS IEDLYAFREE LLVKA THASTIPIVSAVOHE
TOYTLCHFADDVAPPPBAABEIVCKGSEEQVQVFGGYLRHILDHISROLLTSKKQQLLFW
HRFLDRAEPVTTAQQQLDSIETAIQROVQOKITHESQHYDMIGRAUQDDLVGMTCRLOS
LKKMLDGALSHKALDLQVRCHQUKKSLTYPRO JQQAGOKLSHWRQQLDTLISPRLHVÖKE
EYPHKIPHELKHAINVLEQQLRCHVQKLELLGPRLGSCGELNLQNQKIAVANVKETLATIL
ERRYBHDVARYGALKEQDLHSLNARNVLKROYAMLFDFHENDAMISVDSLQENARVRIQLQ
LOGATLTYTNIETCKLIKG

CPn_1063 1219900 1260712

Trase
FCRESMRIKFRENKERKMTR
SPETSLRAIHEVINTTGAFLWUGANNYHELSGAFTJEISLPMLKEVYVEFVLVOHSERR
HIFGESDAFIASKVKSVAQAGLVEVVLCVGESLEVREEGKAHQVIKKQLLLGLECMDNOSE
FLIAYEFWMAIGTGKVAEASDVQDIHMFCREVVAERFSEATAEEISILYGGSVKVDNAQR
FGQCSDVDGLLVGGASLECQSFFEVAKNFNV

CPn_1064 122/716 1220995
The administrate model of the product in the debtatic PMBL is the FIT THE PRODUCT OF T

CPn_1065 (221140 1220928 No robust homolog present in Genebank/EMBL as of 11/7/98 RHRLGRHRRTSDPCFLFFFSIPEESLPPDSCRLNQMPKHEHLPSILLKKPIIDYLKITSI YEKAIFNTGLP

CPn_1066 1221132 1221488
No robust homolog present in Genebank/EMBL as of 11/7/98
SMSLNKEIGMTVLFYAFLFIFFLEVILCGLILVQESKSMGLGSSFGVDSGDSVFGVSTP
DILKKYTSWCAV#FCIGCLLLSFSTNLLGKKLDAKEFLLPAAEESDTQASSESVEADES

CPn_1067 1221675 1222292

def-Polypegtide Deformylase
IQVLVVRDFFTELCQAHVOTMIRRLEYYGSPILRKKSSPIAEITDEIRNLVSDMCDTMEA
HRGVGLAAPØVGKNVSLFYMCVDRETEGGELIFSESPRVFINPVLSDPSETPIIGKEGCL
SIEGLRGEYFRPOKITVTAMDLNGKIFTEHLEGFTARIIMHETDHLNGVLYIDLMEEPKD
PKKFKASLEKIKRRYNTHLSKEELVS

CPn_10d8 1223267 1222365
rnhB-RADORUClease HILL
MSCMPPFFVVTLTTSAONNIRDQLKEKNFIFSQPONTVFQARSNTVTCTLYPSGKLVIQG
KGSEMFIEFFLEPEILHTFTHARVEQDLRPRLGVDESGKGDFFGPLCIAAVYASNAEILK
KLYENKVODSKNIKDTKIASLARIIRSLCVCDVIILYPEKYNELYGKFQNLNTLLAWAH
TVINNLAPKPAGDVFAISDOFAASEYTLLKALQKKETDITLIQKPRAEQDVVVAAASILA
RDJFVQSIQKLEEQYQVQLPKGAGFNVKAAGREIAKQRGKELLAKISKTHFKTFDEICSG

Pn_1069 1223507 1223941
/fga-HTH Transcriptional Regulator
/INDEHIHKEILHIGEIFRSSRESGSLSLKDVEAATSIRYSCLEAIEQGCLGKLISPVYA
QGFIKKYATYIGLDGDSILQEHPYVMKIFKEFSDHNMEMLLDLESMGGRNSPERAIHSWS
NLWWAGLIIIGGIMYWMIGSLFSIF

CPn_1070 1225523 1224144

No robust homolog present in Genebank/EMBL as of 11/7/98

RRSLMTFPCGNCNCYYRETPPPNFGGEDIPLQEGGGSGGGGRVITQOPGTGGREMGISL
GSDNVLGMVEQAGSLLNNLLDSARMQRLGHYCYRTGTPWCREHCPGFLQWIWGGCCACCL
ETVDDPDNPSAQFLQQLIQOYGPICVGMSFQQLPHCTQXIEQGEPLGDGNQEVENGCKL
HRELLKAAQPRCMGESLVKLLQNNGLGEDMQQTPPWSLILQAVSEGALSFVTSSDNPPTC
WILQPEQQPCPPPPTDEEQLGAVGGAPAPQQKKHPAQECRVTCKLNFRTLLQKLSRLEV
LSLESGYKRGLGQAAKQIVDLIKKSLKRLVASDLATFLGPGIGLSLESQVFEVLVLLCLL
SKGYLPLDPLHPEQTVLDPRVQGPWQRILRKVLVTTTAGENIWRQTQGEAPRQAPPPPDP
WDDDEIERDGIVTGGGFGIPCQCLRCWRKLPTEKRPNRWL

CPn_1071 1227336 1225885

No robust homolog present in Genebank/EMBL as of 11/7/98

KGTTMVCPNNSWFRWCGNFNCEWVEVTTTEETTRQSASDISEEAGSSGGAPITTOPTKI

TKVEKRVOFNTAGODESTIHHIQEAGELVDSILSHRRTQGCTEYCYDSYATGCGQRCGSF

GRLIGGTYKACCLDREDNQVAGLVHECEQTHGPIAVALAAKTMGLNLMELVEKNTILSEE

QKNEFRQHCSEAKTQLYGTMQSLSQNFFLEGVNSIRERGLDDSLVQAVLSFIATRSWEKT

IESEEASGTSSASNSTRIPACYILNTSPLTTSRLSCGSRDARRPSSVGAEPQYVAKKYN

NGMARQLGKIQVTNLKYTGDFSALGPFGLLIVKMLNSFLLSASQSTSSILKHTGGEICYTC

PNFRDIVVLLMLAIGYCPANTDETSVVDIHHIDDPIMTIFYRLQYSYRTGKTSASFLKKK

PSLVRQESLDCPTPAESVPLMSSLEEEDENEDDDEDGNLAYQQRILECSGHLQTLFLGIK

INKE

CPn_1072 1227924 1228835

No robust homolog present in Genebank/EMBL as of 11/7/98

KKDYILHANWCCWKQMLKIQKKRMCVSVVITVGALVGFFNSADAAPKKKKIPIOILYSFT

KVSSYLKNEDASTIFCVDVDRGLLQHRYLGSPGWOETRRRQLFKSLENQSYGNERLGEET

LAIDIFRNKECLESEIFEGMEAILANSSALVLGISSFGITGIPATLHSLLRQNLSFQKRS

IASESFLLKIDSAPSDASVFYKGVLFRGETAIVDALSQLFAQLDLSPKKIIFLGEDPEVV

QAVGSACIGWGMNFLGLVYYPAQESLFSYVHPYSTATELQEAQGLQVISDEVAQLTLNAL

CPn_1073 1229011 1229832
Predicted OMP (CT371)
MRRYLFMVLALCLYRAAPLEAVVIKITDAQAVLKFAREKTLVCFNIEDTVVFPKQMVGQS
AWLYNRELDLKTTLSSEQAREQAFLEMMGISFLVDYELVSANLRNVLTGLSLKRSWVLGI
SQRPVHLIKNTLRILRSFNIDFTSCPAICEDOWLSHPTKDTTFDQAMAIEKNILFVGSLK
NGQPMDAALEVLLSGISSPPSQIIYVQDAERLRSIGAFCKKANIYFIGMLYTPAKQRVE
SYMPKLTAIQWSQIRKNLSDEYYESLLSYVKSK

RNA SECTION

LIMENA 13849 (18074

Ribonus Isaas P. RNA 607342 697549

ERNA	Begin 	End 89728 91070 91070 296147 296224 409922 462214 672318 677337 739486 781980 84896 784991 843999 1030533 999949 961536 807341 786708 631373 6269070 268992 164318 87450	Type- Thr Med Asp Arg Leu Asp Clus Arg Leu Lyla Arg Clus Arg Clus Arg Clus Arg Clus Arg Clus Arg Clus Arg Clus Arg Clus Arg Clus Arg Clus Arg Clus Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg	CODOT COTA COTA COTA COTA COTA COTA COTA CO